2-1-2020

DE-CODING THE IMPACT OF EVOLVED CHANGES IN GENE EXPRESSION AND CELLULAR PHENOTYPE ON PRIMATE EVOLUTION

Trisha Zintel
University of Massachusetts Amherst

DE-CODING THE IMPACT OF EVOLVED CHANGES IN GENE EXPRESSION AND CELLULAR PHENOTYPE ON PRIMATE EVOLUTION

A Dissertation Presented

By

TRISHA M. ZINTEL

Submitted to the Graduate School of the

University of Massachusetts Amherst in partial fulfillment

of the requirements for the degree of

DOCTOR OF PHILOSOPHY

February 2020

Program in Molecular and Cellular Biology

DE-CODING THE IMPACT OF EVOLVED CHANGES IN GENE EXPRESSION AND CELLULAR PHENOTYPE ON PRIMATE EVOLUTION

A Dissertation Presented

By

TRISHA M. ZINTEL

Approved as to style and content by:

Courtney C. Babbitt, Chair

Elena Vazey, Member

Jason M. Kamilar, Member

Patricia Wadsworth, Member

Scott S. Garman, Graduate Program Director
Graduate Program in Molecular and Cellular Biology

DEDICATION

To my parents, Brian and Donna Zintel, for their unrelenting support.

ACKNOWLEDGEMENTS

Thank you Courtney for your continued support and patience throughout this process. I am so grateful to have had the opportunity to learn from you.

Thank you Jason Kamilar for your valuable insight and for facilitating my transition into biological anthropology.

Thank you Elena and Pat for your providing me guidance, support, and extensive expertise.

Thank you to Jason Pizzollo for the years of listening to my ideas, good and bad, and for being there throughout this journey. Thank you to Christopher Claypool, as well as the current and past undergraduate members of the Babbitt lab – it's been a pleasure to work with you.

Thank you to Sheryl Smith, for your fierce belief in my abilities and continued support after I left your lab that has been critical throughout my time as a graduate student. Your gracious, supportive attitude is mirrored in your other previous students, Emily Ng and Amber Weiner, who I cannot thank enough for being lasting pillars of support.

Developing a feeling of community and camaraderie has been critical for my professional development as a scientist as well as my personal growth, and for that I have to thank all of the friends I have made throughout my time here. It (apparently)

took a (large) village to support me throughout my experience as a graduate student, and I am so very grateful to everyone – I could not have done it without you. I am not going to attempt to list everyone here, but your friendship has been instrumental to my success. Thank you to those I met early on in graduate school who made the transition to Amherst easier than I anticipated (Ira Male, Katherine Zabala Sanidad, Edwin Murenzi, Korin Albert). Thank you to the OEB students I somehow ended up becoming so close to and for accepting me graciously with no (or very few) questions asked (Lauren Alteio, Laura Hancock, Kit Straley). Thank you to the graduate students I had the opportunity to mentor – you've become incredible friends and I am so lucky that the mentoring ended up being bidirectional (Olivia Williamson, Noelle Dziedzic, Heather Sherman). Finally, thank you to the Kamilar lab members that I owe so much professional and personal support to (Anna Weyher, Mandy Fuchs, Andrew Zamora), and to everyone else I've met along the way (Dan, Zander, Chris, Lisa, Kevin, and many others, local and distant).

Last, but far from least, I want to thank my family. Despite not really having any idea why I was "still in college", their patience and unwavering support has been invaluable. Specifically, I want to thank my parents for lending me their confidence in my abilities when I could not find it in myself. Thank you for always believing in me and supporting all the big ideas and aspirations I had. I could not have done this without you.

ABSTRACT

DE-CODING THE IMPACT OF EVOLVED CHANGES IN GENE EXPRESSION AND CELLULAR PHENOTYPE ON PRIMATE EVOLUTION

FEBRUARY 2020

TRISHA M. ZINTEL

B.A., ARCADIA UNIVERSITY

Ph.D., UNIVERSITY OF MASSACHUSETTS AMHERST

Directed by: Professor Courtney C. Babbitt, PhD

The goal of the dissertation work outlined here was to investigate the influence of proximal processes contributing to evolutionary differences in phenotypes among primate species. There are numerous previous comparative analyses of gene expression between primate brain regions. However, primate brain tissue samples are relatively rare, and my results have contributed to the pre-existing data on more well-studied primates (i.e. humans, chimpanzees, macaques, marmosets) as well as produced information on more rarely-studied primates (i.e. patas monkey, siamang, spider monkey). Additionally, the primary visual cortex has not previously been as extensively studied at the level of gene expression as other brain regions in primates. My investigations of differences in cell biology between human and chimpanzee fibroblasts and iPSC-derived neural cells will contribute to the fields' understanding of the influence of gene expression on differences in cell biology. While iPSC technology has been used extensively to investigate neurological disease *in vitro*, it has not been used to investigate differences in neural function between species. These data will be relevant

both for determining proximate influences on evolutionary differences in neural function across primates and the limitations of use of non-human primate models of neurological disease.

TABLE OF CONTENTS

Page

LIST OF TABLES

Table Page

LIST OF FIGURES

Figure Page

INTRODUCTION

The importance of investigating primate variation to better understand uniquely-human phenotypes

Though primates exhibit widespread variation in many phenotypes, including anatomy, behavior, and cognition, the extent of these phenotypic differences is not substantially larger than differences in genome sequence (Varki and Altheide, 2005; Varki and Varki, 2015). In 1975, decades prior to sequencing the entire human genome, King and Wilson hypothesized that phenotypic differences amongst humans and chimpanzees must largely be due to changes in gene expression, as coding sequence differences were not nearly divergent enough to account for those differences (King and Wilson, 1975). Subsequent investigations into differential gene expression across species, and primates in particular, have reinforced the hypothesis that non-coding, *cis*-regulatory changes are critical for these large differences in phenotype (Haygood *et al.*, 2007; Babbitt *et al.*, 2010; Bauernfeind *et al.*, 2015). Recent work has determined that changes in gene expression in the brain have been found to be both functionally and evolutionarily important (Khaitovich *et al.*, 2004; Oldham *et al.*, 2006; Babbitt *et al.*, 2010, 2011; Kuzawa *et al.*, 2014; Bauernfeind *et al.*, 2015). Changes in non-coding regions of the genome identified to be under positive selection in humans are enriched for biological processes and molecular functions that are neural-specific (e.g. synaptic signaling) and those related to cellular metabolism (e.g. carbohydrate and lipid metabolism) in various studies of human, chimpanzee, and rhesus macaque brains (Babbitt *et al.*, 2010; Haygood *et al.*, 2010; Bauernfeind *et al.*, 2015).

Within primates, selective differences in the genome can be linked to diet and metabolism, suggesting selection has optimized different metabolic processes in lineage-dependent ways (Stringer & Andrews, 1988; Schaffner *et al.*, 2005; Fagundes *et al.*, 2007;

Babbitt *et al.*, 2010, 2011; Haygood *et al.*, 2010; Bauernfeind *et al.*, 2015). The human brain is more energetically costly than that of other primates, utilizing ~ 20% of all of the body's metabolic resources, in comparison to non-human primate brains that use less than 10% (Mink *et al.*, 1981; Hofman, 1983). Importantly, allometry alone does not explain the increase in human brain appropriation of glucose metabolism at this proportion (Karbowski, 2007; Martin, 1981; Yu et al., 2014). There is evidence that sheer increase in neuron number can explain at least part of the energetic demand of the human brain (Herculano-Houzel, 2011). However, interspecies differences in the contribution of metabolism of various neural cell types to metabolic capacity at the organ level remain largely un-explored.

Many of these changes in brain metabolism have been hypothesized to coincide with other trait changes, particularly those related to shifts in diet known to be important in hominin evolution, such as an increase in meat products, increased quality of food, and agriculture (Aiello and Wheeler, 1995; reviewed in Babbitt et al., 2011; Brown et al., 1985; McHenry, 1994, 1992; Peters, 2007; Shea, 2007). A recent study investigating an extensive number of primates is consistent with this notion, determining that the best predictor of brain size was diet, rather that social group structure (DeCasien et al., 2017). The expensive-tissue hypothesis posits that a trade-off in energy allocation for the development of a larger, metabolically demanding brain in primates coincided with a reduction in gut tissue (Stearns, 1992; Aiello & Wheeler, 1995; West *et al.*, 2001; Pontzer *et al.*, 2014). Similarly, an increase in energy-storing tissue (adipose tissue) at the expense of energetically demanding muscle tissue may have also allowed for greater allocation to a larger brain (Leonard et al., 2003; Leonard and Marcia, 1994; Leonard and Robertson, 1997). Investigations of energy budget are a part of a larger effort to elucidate trade-offs between metabolically relevant tissue types throughout primate evolution that might explain unique primate traits, such as longer lifespans, protracted development, and increased relative size of metabolically expensive tissues (i.e. brain) (Charnov and Berrigan, 1993; Lovegrove, 2009; Pontzer et al., 2014; Snodgrass et al., 1999; Speakman, 2005). Indeed, there is

evidence that the higher metabolic costs of the human brain influences the protracted development of body growth rate (Kuzawa et al., 2014). Evolutionary differences in brain metabolism are a subset of studied differences in metabolic traits that exhibit intriguing differences across primate species. Primates exhibit a lower total energy expenditure (TEE) to body size ratio than non-primates, and furthermore, humans have greater TEE than closely related great ape species (chimpanzees, bonobos, gorillas, and orangutans) due in large part to increased basal metabolic rate (BMR), all of which is consistent with the reallocation of metabolic investment to brain (Pontzer et al., 2016, 2014). These *in vivo* (whole organism) studies further suggest an important link between evolutionary differences in metabolism and the uniqueness of the primate brain.

Similarly to organism-level investigations, there is also molecular evidence supporting the evolution of metabolic processes (e.g. oxidative phosphorylation) in the primate brain with phylogenetic proximity to humans. Within anthropoids, genes encoding the subunits of cytochrome c oxidase, the final component of the electron transport chain, show an accelerated rate of evolution in their sequences compared with any other placental mammals [1]. These changes at the gene level are indicative of increased control over the mechanisms that process glucose [2-6]. Further understanding gene-phenotype relationships between genetic changes (both in coding and non-coding regulatory portions of the genome) and observed metabolic differences in primates will contribute to a greater understanding of proximate influences on larger evolutionary trends in primates.

Evolved changes in gene expression have significantly impacted primate evolution in a variety of traits

Cellular metabolism involves the breakdown of fuel molecules to produce energy or other molecules through multiple interconnected pathways, including glycolysis, oxidative

phosphorylation, and the pentose phosphate pathway. Enrichments for metabolic processes in genes and gene regulatory regions undergoing positive selection is a common thread in gene expression analyses from whole primate brain tissue (Babbitt et al., 2010; Bauernfeind et al., 2015; Haygood et al., 2010, 2007; Kosiol et al., 2008; Uddin et al., 2008). However, there are lineage-dependent differences in the specific pathways enriched between species (e.g. glucose and carbohydrate metabolism in humans and glycogen and acyl-CoA metabolism in chimpanzees) (Haygood et al., 2007; Kosiol et al., 2008; Uddin et al., 2008) as well regions of accelerated genomic change in a lineage-dependent manner (Pollard et al., 2006; Prabhakar et al., 2006). These findings highlight a need to investigate not only glucose metabolism and energy production but also that of other macromolecules (e.g. lipids, amino acids, nucleic acids) for a more comprehensive understanding of differences in cellular metabolism in neural cells of primates. Furthermore, unpublished data from our lab from a different study to the one planned here has determined that genes significantly differentially expressed in multiple brain regions between humans and chimpanzees are enriched for metabolic functions related to lipid and carbohydrate metabolism, suggesting this is important for many cortical areas (further discussed in Goal 2: Pilot Results). While previous work at the organ level highlight a difference in brain metabolism, a cell-type specific comparison is necessary to understand distinct cellular contributions to interspecific differences in neurological function (Gallego Romero et al., 2015). Recent advancements in generating induced pluripotent stem cells (iPSCs) and further development and refinement of protocols for differentiation of iPSCs has revolutionized our ability to investigate uncommon cell types without invasive collection techniques. These advancements make investigating evolved differences in cell biology in primates, for which samples are rare, far more feasible. Comparisons of gene expression between iPSCs of different species have found them to be a useful tool for studying inter-species differences, including those between closely related primates (Gallego Romero et al., 2015; Wunderlich et al., 2014).

Harnessing contemporary cell and molecular techniques to investigate classic questions in biological anthropology

Evolved differences in metabolic investment may be the basis for a number of primate-specific phenotypes, including those that are uniquely-human, such as slow reproduction and growth and correspondingly longer lifespan than other placental mammals (Charnov and Berrigan, 1993; Pontzer et al., 2014; Snodgrass et al., 1999). Adaptation can act on metabolic phenotypes through alterations to energy budget including reductions of or increases in total energy budget or differential allocation of energy within energy budget (Lovegrove, 2009; Pontzer et al., 2014; Speakman, 2005). In addition to energy allocation, adaptive changes in diet (e.g. ability to differentiate nutrient poor vs nutrient rich food, tool use) also likely contributed to the enlargement of the brain, particularly in the human lineage (reviewed in Babbitt et al., 2011). The proposed work aims to bridge hypotheses about the influence of gene expression on phenotypic differences amongst primates with trade-off hypotheses about the evolution of a larger brain at the expense of other metabolically demanding tissue types (King & Wilson, 1975; Leonard & Marcia, 1994; Aiello & Wheeler, 1995; Leonard & Robertson, 1997; Leonard *et al.*, 2003) by investigating altered gene expression in an under-studied, but critical brain cell type with known metabolic relevance. My results would provide valuable insights into astrocyte-mediated differences in metabolic brain function as a proximate mechanism by which the ultimate evolutionary trajectory of human brain evolution may have occurred. The novel technical approach I outline here would provide the larger biological anthropology community with candidates of specific metabolic processes for which primates are experiencing adaptation, and have the potential to inform studies of metabolic phenotypes in primates that are not astrocyte- or brain-specific (i.e. evolved differences in primate energy budget). Determining processes in which there are lineage-specific differences in gene expression and cellular function, as I aim to do in with this project, can indirectly inform studies of organism-level

energetics by elucidating what processes, and the functional consequences of, are experiencing adaptation on the human lineage.

Summary of presented work

In conducting my dissertation research, I investigated how differences in gene expression influence phenotypically diverse traits between primate species. To do this, I coupled next-generation sequencing (RNA-Seq) and cell biology techniques (cell culture, induced pluripotent stem cells (iPSCs), immunohistochemistry) to investigate the influence of differential gene expression on observable differences in tissue- and cellular-level traits between primate species. This included three major projects: 1) investigating the evolution of gene expression in the primary visual cortex of primate brains, 2) determining cellular differences in adhesion and migration between humans and chimpanzee fibroblasts, and 3) investigating evolutionary differences between neural cells in different primate species using iPSC-derived neurons and astrocytes. Each of these projects are in different stages of completion. Here, I briefly introduce each but full descriptions of these projects follow this introduction.

Chapter 1: Ecological trait differences are associated with gene expression in the primary visual cortex of primates. There is evidence that divergence in substructures of the brain may be equally as important phenotypically and evolutionarily as relative brain size (Aristide *et al.*, 2016). Furthermore, multiple studies on brain gene expression have identified differences in gene expression based upon brain region (Khaitovich *et al.*, 2004; Oldham *et al.*, 2006; Bakken *et al.*, 2015; Bauernfeind *et al.*, 2015). The goal of this project was to determine differential expression across a wide range of primates in an understudied brain region, the primary visual cortex (V1). This project is a subset

of a larger project comparing global gene expression (via RNA-Seq) across four brain regions (cerebellum, hippocampus, prefrontal cortex, and visual cortex) between 18 primate species, of which I am a co-author on and played a main role in sample preparation, sequencing, and computational pipeline development. The data presented here for this project focuses on differential gene expression in the V1 using genome-guided determination of gene expression in 13/18 primate species for which genomes are available. Significantly differentially expressed genes were determined between species (i.e. human – chimpanzee), between groups of species belonging to different taxa (i.e. all apes vs all old-world monkeys), and between groups of species that differ in phenotypic traits (i.e. all arboreal species vs all terrestrial species). Differential expression (DE) of genes in the visual cortex between species were enriched for many different pathways, including those related to disease, adhesion, signaling, metabolism and neural-specific processes. Genes DE between humans and chimpanzees were enriched for more unique pathways than any other interspecies comparison. Differential expression detected based entirely on comparing arboreal to terrestrial species was enriched for synaptic neural processes more than any other DE comparison based on other phenotypes (i.e. color vision) or differences in phylogeny. I found that while metabolic signaling differs consistently across the primate tree, neural-specific signaling (e.g. synaptic signaling) is far more conserved, except with evolutionary proximity to humans: humans and chimpanzees were the only two species that exhibited significant differences in synaptic signaling. This project is completed and a manuscript of these results is in progress to be submitted for publication in November 2019 to BMC Evolutionary Biology.

Chapter 2: Chimpanzee fibroblasts exhibit greater adherence and migratory phenotypes than human fibroblasts. Primates differ in a number of different phenotypic traits, including that of disease (Varki & Altheide, 2005; Varki & Varki, 2015). More specifically, it has been noted that cancer rates differ significantly between humans and chimpanzees. Epithelial neoplasms, such as breast or lung carcinomas, are responsible for at least 20% of human deaths but account for less that 2% of chimpanzee deaths (Puente *et al.*, 2006). Previous work from our lab determined that human and chimpanzee fibroblasts are significantly different in (A) adhesion-related gene expression when mimicking a wound healing / cancer response *in vitro* and (B) proportions of focal adhesions *in vitro* (Advani *et al.*, 2016).The goal of this project was to determine if these previous findings related to adhesion translates into fundamental differences in the cellular phenotypes of adhesion and migration between human and chimpanzee fibroblasts *in vitro*. In collaboration with Shelly Peyton's lab at UMass Amherst, preliminary experiments comparing human and chimpanzee fibroblast adhesion and migration in the presence of external fluid force determined that chimpanzee fibroblasts adhered at greater proportion within one hour than human fibroblasts did without observable differences in migration speed or distance over the entire 12-hour time course. A wound healing assay determined that chimpanzee fibroblasts exhibit greater adhesion and migration than human fibroblasts. Finally, the same serum-challenge that mimics cancer progression / wound healing that produced the aforementioned previous results of differences in adhesion gene expression will be replicated in these experiments to determine interspecies differences in adhesion and migration in this context. This project is completed and a manuscript of these results is

in progress to be submitted for publication in November 2019 to Evolutionary Medicine and Public Heath.

Chapter 3: Investigating cell-type specific shifts in metabolic gene expression between human and chimpanzees. Project one focused on whole-brain tissue differential expression across primate species and determined enrichment for metabolic and neural-specific processes in differential expression between humans and chimpanzees. However, the brain consists of several different cell types for which there are cell-type specific metabolic and neural functions. For example, astrocytes promote neural function by buffering extracellular conditions and by shuttling important metabolites to neurons and even play an important role in neuron-neuron communication (i.e. synaptic transmission) (Almad *et al.*, 2016; Chandrasekaran *et al.*, 2016; Zhao *et al.*, 2017). To investigate differences in species-specific gene expression and function in a cell-type specific manner, I compared human and chimpanzee iPSC-derived neurons and astrocytes. This involved differentiation of iPSC lines into multipotent neural progenitor cells (NPCs) and subsequent differentiation of separate populations of NPCs into neurons and astrocytes. Homogenous populations of neurons and astrocytes, per each species, will be validated via immunohistochemistry for cell-identity and will be used for RNA-Seq to determine differential gene expression in a cell-type by species manner. Our results comparing expression between species demonstrates that astrocytes play a significant role in altered metabolic gene expression.

With these outlined projects, I have further elucidated the influence of proximal processes contributing to evolutionary differences in phenotypes among primate

species. There are numerous previous comparative analyses of gene expression between primate brain regions. However, primate brain tissue samples are relatively rare, and my results have contributed to the pre-existing data on more well-studied primates (i.e. humans, chimpanzees, macaques, marmosets) as well as also produced information on more rarely-studied primates (i.e. patas monkey, gibbon, spider monkey). The primary visual cortex has also not been as extensively studied as other brain regions in primates. My investigations of differences in cell biology between human and chimpanzee fibroblasts and iPSC-derived neural cells will contribute to our understanding of the influence of gene expression on differences in cell biology. While iPSC technology has been used extensively to investigate neurological-disease *in vitro*, it has not been used to investigate differences in neural function between species. These data will be relevant both for determining proximate influences on evolutionary differences in neural function across primates and limitations of use of non-human primate models of neurological disease.

CHAPTER 1

ECOLOGICAL TRAIT DIFFERENCES ARE ASSOCIATED WITH GENE EXPRESSION IN THE PRIMARY VISUAL CORTEX OF PRIMATES

1.1 Abstract[1]

Background**:** Primate species differ drastically from most other mammals in how they visually perceive their environments, which is important for foraging, predator avoidance, and detection of social cues. Although it is well established that primates display diversity in color vision and various ecological specializations, it is not understood how visual system characteristics and ecological adaptations may be associated with gene expression levels within the primary visual cortex (V1).

Results**:** We performed RNA-Seq on V1 tissue samples from 28 individuals, representing 13 species of anthropoid primates, including hominoids, cercopithecoids, and platyrrhines. We explored trait-dependent differential expression (DE) by contrasting species with differing visual system phenotypes and ecological traits. Between 4-25% of genes were determined to be differentially expressed in primates that varied in type of color vision (trichromatic or polymorphic di/trichromatic), habitat use (arboreal or terrestrial), group size (large or small), and primary diet (frugivorous, folivorous, or omnivorous). Our DE analyses revealed that humans and chimpanzees showed the most marked differences between any two species, despite the fact that they are only separated by 6-8 million years of independent evolution.

[1] This chapter is formatted for publication in BMC Evolutionary Biology with a planned submission date sometime in November 2019. The full author list for this manuscript is: Trisha M. Zintel, John J. Ely, Mary Ann Raghanti, William D. Hopkins, Patrick R. Hof, Chet C. Sherwood, Jason M. Kamilar, Amy L. Bauernfeind, and Courtney C. Babbitt

Pathway enrichment analyses of DE genes demonstrated that changes in cellular metabolic pathways (e.g. glycolysis) contribute to altered gene expression in primate V1 more than neuron-specific processes (e.g. synaptic signaling). The exception to this trend is between human and chimpanzee, which exhibited DE for a number of processes related to cholinergic and GABAergic synaptic signaling.

Conclusion: Our data significantly expand the number of primate species for which V1 expression data exists. These results show a combination of species-specific and trait-dependent differences in the evolution of gene expression in primate V1. We also show that human-specific changes in brain gene expression extend to the primary visual cortex in a manner similar to that reported of other brain regions.

1.2 Background

In most mammals, vision is critically important for the purposes of foraging, predator avoidance, and mate recognition [7-11]. However, primates are distinguished from other mammals by unique visual traits, including high visual acuity and variation in color perception [8, 11-15]. Primates with forward-facing eyes have high orbital convergence and correspondingly enhanced binocular vision resulting in greater depth perception (stereoscopic vision) and increased visual acuity [16]. Primates also possess a uniquely specialized fovea, a localization of cone receptors in the retina, that allows for greater resolution [17, 18]. Among primates, visual acuity is highest in diurnal haplorrhine species [19-23]. Trichromatic color vision is a hallmark of hominoid and cercopithecoid species, but color vision varies significantly across platyrrhine species as well as within some species that have polymorphic sex-linked di/trichromacy [24].

Higher visual acuity and a shift to trichromatic color vision are hypothesized to confer a number of benefits across primate species. High visual acuity has been suggested to be adaptively favorable in relation to diet (i.e. detecting insects) as well as for arboreal species

navigating tree limbs [16]. At the level of opsin gene conservation across primates and tree shrews, typical light conditions that require enhanced visual acuity are the same as those in which color vision is favored, representing a putative adaptive link between enhanced acuity and trichromacy [14]. The evolution of a third opsin gene increases the range of detectable wavelengths (especially in the range of red hues) and has been hypothesized to be influenced by foraging and social pressures [25, 26]. The foraging hypothesis stresses the importance of detecting red ripe fruits and young leaves against a green foliage background [7-9, 27, 28]. The social signal hypothesis states that wavelength sensitivity for red hues is important for accurate interpretation of skin color changes indicative of receptivity, sociality, or health [25, 29, 30]. Evidence in a lemur species with mixed populations of trichromats and dichromats showed that there is a group level benefit when at least a single trichromatic individual is present, likely linked to the ability to locate ripe fruit during the dry season, thus positively influencing fitness [28].

The primary visual cortex (V1) is the first cortical region in the visual pathway and where visual inputs converge from both eyes. Inputs from both eyes that impart binocular vision are critical for detecting edge orientation, thus playing an important role in depth perception [15, 31]. Increased orbital convergence allowing for binocular vision correlates with an increase in size in brain regions associated with visual perception [32]. Neuronal density within V1 of primates is more than twice that of other mammals, reflecting the importance of vision for primates [15, 16]. There is also known variation in V1 neuronal cytoarchitecture across hominoid species, including molecular differences in cytochrome oxidase staining in layer 4A [33, 34] and interneuron density [35, 36] as measured by immunohistochemistry. Taken together, these findings may indicate an adaptive advantage of visual processes that could differentially influence V1 function in primates. However, we still have little knowledge about the genetic mechanisms driving V1 variation across species and how this correlates with differences in visual acuity and perception.

King and Wilson [37] hypothesized that the extensive phenotypic differences between humans and chimpanzees must be due to changes in gene expression rather than coding-sequence evolution, given the high degree of similarity at the nucleotide and protein levels. Investigations of differential expression (DE) across primates have demonstrated that *cis*-regulatory changes are critical for large-scale differences in phenotype [38-42]. Gene expression tends to vary in different tissues of the body, and different regions of the brain are known to have region-specific profiles indicative of localized regulation [43-48]. Enrichments for neuron-specific processes (e.g. synaptic signaling) and metabolism (e.g. carbohydrate and lipid metabolism) have been reported in human brain gene expression in comparison to that of chimpanzee and rhesus macaque [39-41] and may be indicative of region-specific differences in neural function.

There have been few investigations of interspecific gene expression differences of the visual cortex of primates [44, 49-51]. Khaitovich et al. [44] investigated a number of brain regions, including V1, with the primary focus on determining the extent of region-specific differences in expression between 3 primate species (human, chimpanzee, and macaque). These authors found that regions of the cerebral cortex, including V1, differ significantly from other brain regions, such as the cerebellum, and that the degree of intraspecific variation in expression between these brain regions was significantly greater in human than chimpanzee. Furthermore, they reported that in the cerebral cortex of both species, V1 was an outlier among other regions investigated (dorsolateral prefrontal cortex, anterior cingulate cortex, and Broca's area) [44]. Another study found coexpression networks specific to V1 in both humans and chimpanzees and found the greatest degree of interspecific conservation in V1 compared to any other region investigated [51]. The authors attributed these findings to the sensory nature of V1's function for which humans and chimpanzees presumably differ very little [51]. A comparison of transcription between human, macaque, and mouse V1 using microarrays found that primate (human and macaque) V1 expression was conserved in comparison to mouse, suggesting that

the conserved genomic profiles of V1 extends beyond primates [49]. These studies all represent important initial investigations of V1 gene expression, however, they largely investigate only a few primate species (humans, chimpanzees, rhesus macaques) and emphasize other brain regions over V1.

To understand primate visual system adaptations better at the gene expression level, we used RNA-Seq to quantify gene expression in V1 of 13 primate species across three major clades (hominoids, cercopithecoids, and platyrrhines), significantly increasing the number of primates for which V1 gene expression data exists. This broad phylogenetic sampling of phenotypically diverse primates allowed us to explore how differences in gene expression were associated with variation in traits that contribute to a species' visual perception, including differences in species-typical color vision (trichromats vs polymorphic di/trichromats), and ecological variables such as habitat use (arboreal vs terrestrial), average group size (large vs small), and typical diet (folivore vs frugivore vs omnivore). We also evaluated the influence of sex on gene expression in V1. We found significantly differentially expressed genes in V1 when comparing global gene expression based upon extremes in species-typical traits for color vision, habitat use, group size and diet, but not sex. Pathway enrichment analyses of genes exhibiting differential expression (DE) show that these expression changes correlate with variation in visually-relevant traits, as well as expression differences between species, are largely driven by altered metabolic signaling. Human and chimpanzee V1 expression is notably different from other species, and though they do not differ in the visual ecological traits investigated here, they are divergent from one another in important neurological signaling processes.

1.3 Methods

1.3.1 Samples and Library Preparation

Tissue samples were collected from 28 individuals representing 13 primate species (n = 1-3 per species; Figure 1, SI Table 1.1). Frozen brain samples from captive adult primates free

of known neurological disorders were obtained from various research institutions and zoos. All individuals had been cared for according to Federal and Institutional Animal Care and Use guidelines and died of natural causes. All tissue was collected and stored at -80°C with postmortem intervals of less than 8 hours. Further details about the samples can be found in SI Table 1. Due to the opportunistic nature of sampling such a phylogenetically broad group of primates, there is unavoidable variation in age and sex of some individuals sampled (SI Table 1.1).

Each sample was dissected from the cortex of the medial aspect of the occipital pole, surrounding the calcarine sulcus. The thin, striate cortex of V1 was visually identified in each sample. All dissections included the grey matter of V1, extending a small extent into the underlying white matter. Tissue samples were homogenized using a TissueLyser (Qiagen) prior to total RNA extraction using an RNeasy Plus Mini Kit (Qiagen), including a DNase step to remove residual DNA. Total RNA was analyzed for quality using the Agilent Bioanalyzer system (Agilent RNA 6000 Nano kit). RNA integrity varied among our samples due sampling from deceased primates, but there was no bias for species. Using the NEBNext Poly(A) Magnetic mRNA Isolation Kit (NEB), mRNA was isolated from intact total RNA, and cDNA libraries were made from each sample using the NEBNext RNA Library Prep Kit for Illumina (NEB). Library quality was assessed using the Agilent DNA 1000 kit. Pooled samples were sequenced using the Illumina NextSeq 500 platform to produce 75 base pair reads, yielding a minimum of 20 million reads per sample.

1.3.2 Read Mapping and Quantification

Quality-filtered reads were aligned to available primate genomes with Bowtie2 using default parameters for gapped alignments [52]. Current species-specific ENSEMBL genomes were used (*Homo sapiens,* GRCh38.p10; *Pan troglodytes,* CHIMP2.1.4; *Gorilla gorilla,* gorGor3.1; *Pongo abellii,* PPYG2; *Papio anubis,* PapAnu2.0*; Nomascus leucogenys,* Nleu1.0; *Macaca mulatta,* Mmul_8.0.1; *Callithrix jacchus,* C_jacchus3.2.1) [53, 54]. For species for

which there is no publicly available reference genome (*Pongo pygmaeus, Symphalangus syndactylus*, *Macaca nemestrina*, *Erythrocebus patas*, *Saimiri sciureus, Pithecia pithecia,* and *Ateles fusciceps*), reads were mapped to the closest related primate for which there was a genome available. Specifically, *M. nemestrina* and *E. patas* reads were mapped to *M. mulatta, S. syndactylus* to *N. leucogenys*, *A. fusciceps*, *S. sciureus*, and *P. pithecia* to *C. jacchus*, and *P. pygmaeus* to *P. abellii*. Mapping percentages were ≥ 80% using default "—local" Bowtie2 parameters. The "—very-sensitive-local" parameter was used to increase accuracy and alignment percentage of the samples with the lowest mapping percentages (< than 85%; 2 *A. fusciceps,* 1 *S. syndactylus*, and 1 *S. sciureus* samples), and all increased to ≥ 83.5% (see SI Table 1.1). HT-Seq [55] was used to quantify counts per gene for each sample, using ENSEMBL gene transfer files (GTFs) corresponding to the same genome build used for alignment [56]. For each species, homologous genes were matched to the ENSEMBL human reference set of genes using biomaRt [57]. These were subsequently filtered to the 12,564 genes from all species that have high orthology confidence, all of which also had high homolog percent identity to the human query genes as well as gene order conservation scores, as previously determined by ENSEMBL [56]. This resulted in removal of genes from all transcriptomes considered if they did not meet orthology confidence in all species. Finally, we used the R package edgeR [58] to filter out low-expressed genes (counts per million (CPM) > 1 in 1/28 samples), resulting in 12,330 expressed orthologs in our dataset.

1.3.3 Clustering Analyses

We used clustering analyses to determine the variation in our samples. We produced a principal coordinates analysis (PCoA) using the expression profiles of all of the protein coding genes for each of our samples (n = 1-3 per species). To further visualize patterns within our data, we produced phenograms by performing hierarchical clustering of the PCoA distances. The PCoA and dendrograms show that our samples largely cluster by species and clade.

1.3.4 Differential Expression Analysis

Differential expression (DE) analyses were performed for multiple pairwise comparisons using a generalized linear model (GLM) in the R package edgeR (SI Table 1.2) [58]. Normalization of data in EdgeR for DE analyses ensured that DE is not dependent on original tissue size. Gene expression was considered significantly different at a false discovery rate (FDR) of less than 5%. DE between species only included species for which we had more than one sample, including human, chimpanzee, siamang, olive baboon, rhesus macaque, pig-tailed macaque, patas monkey, spider monkey, and marmoset (SI Table 1.2). To determine if there was a relationship between number of genes exhibiting DE between species, we used a Pearson correlation between numbers of genes exhibiting DE and divergence times reported by 10k Trees of the last common branching point of the tree between the two species compared (SI Table 1.2; [59]. The multi-factor GLM allowed us to analyze all data at once to detect DE between species as well as between groups of samples for which the species are known to have distinct differences in visually relevant phenotypes, data for which were collected from the literature [10, 16, 60-63]. These species-typical data were used to label each of our samples for color visual system (trichromats or polymorphic tri/dichromats), average group size (less than 20 individuals per group or greater than 20 individuals per group), primary diet (frugivorous, folivorous, or omnivorous), and primary habitat (arboreal or terrestrial — mapped onto a phylogeny in Figure 1; listed in SI Table 1.1). For each of these phenotype contrasts, all samples of the same species-typical phenotype were grouped together and compared to the group of samples with the opposing species-typical phenotype (e.g. all samples from primarily frugivorous species were grouped together and compared to all samples from primarily omnivorous species for the phenotype-DE comparison frugivorous versus omnivorous diet). DE was also performed on each sample by sex (male or female; SI Table 1.1). Given the comparative, discovery-based nature of our analyses, we did not correct p-values of DE genes for the multiple interspecies or phenotype-DE comparisons, though many would have remained significant.

Our goal was to investigate differences in V1 gene expression between species, but also to attempt to link V1 expression differences with differences in broad, species-level phenotypes. It is important to note that there a number of caveats to this approach relevant to interpretation of results. We used species-typical categories for each of the phenotypes we were interested in correlating with V1 gene expression because we do not have the relevant phenotypic information on each of the individuals from which our brain samples were obtained. Importantly, our samples are from captive living animals that did not necessarily experience the wild, ecological condition typical for their species. Thus, our study design allows us to interpret minimal innate differences in V1 gene expression that may be related to past evolutionary influences from visual ecological pressures, but not from dynamic experience-dependent differences, which would probably yield even stronger signals of ecological correlations. For these reasons, we labeled our samples for phenotype using broad, categories (e.g. primarily arboreal species vs primarily terrestrial species rather than more discrete measurements) and we determined DE by grouping samples by species-typical phenotype and comparing between phenotypically distinct groups. This leads to another important caveat for our phenotype-DE comparisons, which was that our analyses did not allow us to account for phylogeny. We "mapped" these traits onto a phylogeny in Figure 1 to provide transparency about some of the clear influences of phylogeny on our divergent phenotype groups. For example, the investigated hominoid and cercopithecoid species are monomorphic for trichromatic color vision in all individuals regardless of genotype or sex, while the platyrrhine species are polymorphic with trichromatic homozygous females, dichromatic-males or dichromatic-heterozygous females (Figure 1; SI Table 1.1). With these caveats in mind, we were conservative in our interpretation of these analyses.

1.3.5 Categorical Enrichment Analysis of Differentially Expressed Genes

Genes identified as DE for each interspecies and phenotype comparison were subsequently used for pathway enrichment with Kyoto Encyclopedia of Genes and Genomes

(KEGG) [64] to determine differences in signaling and as well as Gene Ontology (GO) cellular component (CC) categorical enrichment analyses [65, 66] to further determine the cellular locality of differences. Enrichments were obtained using EnrichR with *crisp* data sets and filtered by those that have an enrichment p-value less than .05 [67, 68]. To determine the most biologically relevant enrichments, KEGG pathway enrichments were further filtered to include only those with five or more genes per enrichment and GO CC enrichments were further filtered to only include those with ten or more genes per enrichment. Due to the discovery-based nature of our research question and the filtering of categories based upon gene number, we did not restrict our analyses to only categories with significant adjusted p-values for multiple comparisons, of which there were fewer. Similarly, GO CC enrichment terms were also grouped into parent categories related to cellular structures (e.g., "neural cell projection") is a larger category containing GO CC terms such as "axon" and "dendritic branch point" which are encompassed in the established GO CC hierarchy of "cell projections" (SI Table 1.4). The parent category of KEGG "neuron-specific processes" includes enrichments for pathways that are unique to the brain and the majority of these were for synaptic processes (specific neurotransmitters, synaptic activity, and synaptic vesicle function; SI Table 1.3). Metabolic pathways include those related to the metabolic breakdown or synthesis of all major macromolecules, including that of carbohydrates (e.g. "glycolysis/gluconeogenesis"), lipids (e.g., "fatty acid metabolism"), amino acids (e.g. "glycine, serine, and threonine metabolism"), and nucleotides (e.g., "purine metabolism" ; SI Table 1.3). We report all significant categorical enrichments (SI Table 1.3, 1.4), but in our results, largely focus on enrichments for pathways specifically involved in neuronal and metabolic signaling, as previous work has found that these processes exhibit DE and signatures of positive selection in *cis*-regulatory regions when comparing human, chimpanzee, and rhesus macaque brains [39-41] and the importance of metabolism in the brain provides cellular energy and critical synthesis and breakdown of macromolecules [69-73].

1.4 Results

1.4.1 Humans and chimpanzees exhibit the most variation in V1 gene expression

RNA-Seq was conducted on V1 tissue collected from 28 individuals from 13 primate species, including hominoids, cercopithecoids, and platyrrhines (Figure 1, SI Table 1.1; additional details in Methods). Despite the fact that gene expression does not evolve in the same manner as nucleotide sequence [74], due to tissue- and cell-type-specificity within an organism, V1 gene expression profiles overall grouped by species and phylogenetic clade (Figure 2). Interestingly, there was greater variation in V1 transcriptomes within hominoids than cercopithecoids or platyrrhines, predominantly due to distances observed among samples representing human and chimpanzee (Figure 2A). Thus, the expression profiles of other hominoids (orangutan, gorilla, and siamang) were more similar to that of cercopithecoids and platyrrhines than to human and chimpanzee (Figure 2A). Hierarchical clustering of whole transcriptomes revealed similar results, with most samples per species clustering generally according to phylogeny (Figure 2B). The exceptions were for the single representative samples of gorilla, saki monkey, squirrel monkey, and orangutan. Additionally, the two pig-tailed macaque samples clustered away from the other cercopithecoid species. These multivariate analyses indicate conservation in the overall expression profiles of primates with the exception of human and chimpanzee. There are no overt technical factors influencing the out grouping of human, chimpanzee, or pig-tailed macaque samples (e.g. individual sex or age, sample hemisphere of origin, RNA or cDNA library quality, read number, alignment percentages, SI Table 1.1). Our finding of human and chimpanzee divergence from other species is consistent with other studies [38, 41, 47, 48] and highlights that the increased distinctiveness of brain gene expression with proximity to the human lineage extends to V1.

Next, we performed pairwise differential expression (DE) analyses between species to determine the proportion of genes that were statistically significantly differentially expressed

(SI Table 1.2). There was a positive correlation between number of genes exhibiting DE and phylogenetic distance ($r = 0.6$, $p = 0.0001$, Pearson correlation), demonstrating that the number of genes exhibiting DE increases with phylogenetic distance (SI Table 1.2). However, humans exhibit noticeably greater numbers of genes displaying DE when compared to chimpanzee (n= 3,648) and siamang (n= 3,846; SI Table 1.2), despite being relatively closely related. Notably, chimpanzee and siamang do not exhibit as high a degree of difference from one another (n= 2,771; SI Table 1.2). This may suggest that there is significant difference in V1 expression with phylogenetic proximity to humans, and not more generally in the hominoid clade. These results demonstrate that there is a significant degree of DE in V1 across primate species, despite its sensory function.

1.4.2 Metabolic and neuronal signaling are enriched in V1 DE

To determine what biological processes are enriched in genes displaying DE in V1, we performed pathway enrichment analyses for Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway terms of pairwise DE comparisons (Table 2; SI Table 1.3). To obtain an overall view of global processes enriched in primate V1 DE, we grouped specific pathways into broad parent KEGG categories for “signaling”, “immune”, “disease”, “neuron-specific”, “metabolic”, and “other”, generally following the established hierarchy already present in the KEGG ontology (SI Table 1.3; for more details on thresholding and grouping of pathways into parent categories, see Methods). Overall, there was a wide variety of KEGG pathways enriched in these DE comparisons related to inter- and extracellular signaling processes associated with growth and development, as well as more specific pathways related to cellular metabolism (e.g. glycolysis) and neuron-specific signaling (e.g. synaptic signaling; SI Table 1.3). To assess whether metabolic and neural processes routinely differ in the gene expression of primate V1, as well as what interspecies DE comparisons deviated from that trend, we determined what proportion of all the pathways enriched were metabolic or neuron-specific signaling pathways. On average, metabolic pathways accounted for 23.8% of enrichments while neural pathways

only accounted for 7.6% of enrichments (SI Table 1.3). Similarly, enrichments for metabolism were present in all interspecies DE comparisons (34/34 pairwise species comparisons), while enrichments for neuron-specific processes were present in only 23/34 comparisons (SI Table 1.3). Interestingly, there was a reduction in the proportion of metabolic enrichments observed in the human-chimpanzee DE comparison coinciding with an increase in the number of neuron-specific enrichments (Table 1; SI Table 1.3).

Synaptic signaling pathways included in the neuron-specific parent group of enriched pathways included glutamatergic, cholinergic, serotonergic, and GABAergic synaptic signaling as well as "synaptic vesicle cycling" and "long-term potentiation" (SI Table 1.3). Other neuron-specific KEGG pathways that were enriched included axon guidance, neurotrophin signaling, and oxytocin signaling (SI Table 1.3). All eight DE comparisons of chimpanzee to other species had higher proportions of neuron-specific pathway enrichments of any interspecies DE comparison (SI Table 1.3). The DE comparison between human and chimpanzee had the most neuron-specific enrichments of any interspecies DE comparison (n =12; SI Table 1.3), including cholinergic and GABAergic synapse among the top ten most significantly enriched pathways ($p < 0.0001$; Table 1). In contrast, there were no neuron-specific pathways enriched in the top ten KEGG pathways for any other human interspecies DE comparisons (except neuroactive ligand receptor interaction in the DE comparison between human and siamang; Table 1). These results highlight that differences in neuron-specific processes in the V1 are not common across all primates but are more prominent with proximity to human, and most notably between human and chimpanzee.

The V1 region is highly specialized in cytoarchitecture and structure [15], and previous studies have determined that many of the V1-specific differences in gene expression between species are for genes involved in determining this structure [49]. Therefore, in addition to KEGG pathway enrichments, we conducted pathway enrichment analyses for Gene Ontology (GO) cellular components (CC) for pairwise DE comparisons to determine if there were

differences in primate V1 related to specific neuron cellular parts. In many interspecific comparisons, including, but not limited to human vs. chimpanzee, there were several GO CC enrichments related to plasma membrane-associated complexes of known importance for intercellular signaling in the brain, including ATP-coupled ion channels and various neurotransmitter plasma membrane receptors and intermembrane transporters "clathrin-sculpted glutamate transport vesicle membrane", "ionotropic glutamate receptor complex", and "ciliary neurotrophic factor receptor complex" (SI Table 1.4). There were also enrichments for GO CC terms for specialized cell projections that characterize neural cells, including components of dendrites (e.g. dendritic branch point) and types of axons ("C fiber"), as well as for synapse parts and neurofilaments (SI Table 1.4). Interestingly, DE comparisons between human and chimpanzee and other species are the only interspecies DE comparisons to be enriched for dendrite cellular parts and only chimpanzee DE comparisons are enriched for axon cellular parts (SI Table 1.4). Enrichments for GO cellular components corroborate KEGG pathway enrichments and further show that neural signaling complexes critical for neurological signaling and structure, such as plasma membrane associated channels, dendrites, and axons, differ among primate V1.

To determine how metabolic gene expression differed from global gene expression in primate V1, we subset our expression matrix to include only genes involved in metabolic KEGG pathways (n = 1039, Figure 3). When we generated a heatmap where species clustered based upon distance between expression values, we saw that expression profiles of metabolic genes result in clustering of species according to clades: all platyrrhines, hominoids, and cercopithecoids group together with no overt pattern in the metabolic KEGG pathways to which the genes belong (bottom color-coded bar, Figure 3). This demonstrates that while there are significant differences in metabolic gene expression in primate V1, species-specific metabolic patterns still seem to contain substantial phylogenetic signal. To try to elucidate what metabolic processes distinguish each clade, we identified which pathways demonstrated DE between

species of different clades, but not between species belonging to the same clade. Carbohydrate metabolism appears to be the most distinguishing set of metabolic pathways between clades. Hominoids differ from platyrrhines in citrate (TCA) cycle gene expression and differ from both cercopithecoids and platyrrhines for glyoxylate and decarboxylate metabolism (SI Table 1.3). In contrast, platyrrhines differ from both cercopithecoids and hominoids for amino sugar metabolism and glycolysis/gluconeogenesis gene expression (SI Table 1.3). One noticeable exception of a carbohydrate metabolic pathway that differed consistently between species' comparisons of all three cross-clade combinations but never between two species of the same clade was fructose and mannose metabolism, which displayed DE between 13 interspecies comparisons, including six comparisons of cercopithecoids to platyrrhines, five comparisons of hominoids to cercopithecoids, and two comparisons of hominoids to platyrrhines (SI Table 1.3). Taken together, these results show that metabolic changes in primate V1 are common across primate species in a manner consistent with phylogeny and that the metabolic gene expression differences influencing this are primarily related to carbohydrate metabolism. In summary, variation in primate V1 gene expression is not limited to neural signaling processes and in fact seems to be driven to a large extent by altered metabolic signaling, although neural signaling appears to contribute more to differences in expression with proximity to humans.

1.4.3 Visual cortex DE is correlated with differences in color vision, group size, diet, and habitat use

We leveraged the wealth of phenotypic and behavioral data available for the sampled primate species to help understand correlations between V1 gene expression and variation in species-typical color vision, habitat use, diet, group size, and individual differences in sex (summarized in Figure 1; details in Methods and SI Table 1.1) [10, 16, 60-63]. These phenotype-DE comparisons were conducted by grouping all samples according to the species-typical state of their color visual system (trichromats or polymorphic di/trichromats), habitat use

(primarily arboreal or terrestrial), generalized diet (primarily frugivorous, folivorous, or omnivorous), and group size (primarily living in small or large groups) and determining genes exhibiting DE between the phenotype-extremes. We reasoned that this would be a fairly conservative estimate of differential expression because the grouping of such disparate species likely introduces far more variability ("noise") than when comparing two discrete species and results in far fewer numbers of genes exhibiting DE in these phenotype comparisons. Of these trait-based DE comparisons, we detected DE based upon differences in color visual system, primary habitat use, group living size, and primary diet (frugivore-folivore, folivore-omnivore, frugivore-omnivore) but not sex in the samples used for this study (Table 2, SI Table 1.2). The highest proportion of trait-based DE was for differences in color vision (25.3%) and the lowest was for the diet-based comparisons (4.3-7.9%; Table 2, SI Table 1.2). While many of the parent categories of pathways were enriched in all phenotype-DE comparisons, there were unique enrichments for specific pathways depending on the phenotype compared (Table 3). It is important to note that for some traits investigated here that there is a clear influence of phylogeny on differences in phenotype based purely on the species for which we have samples (further discussed in Methods). Because we cannot parse out how much this influences the phenotype-DE comparisons, we are conservative in our interpretations of these analyses.

Color vision has long been hypothesized to be adaptive in primates with numerous complementary and competing hypotheses about the pressures influencing its evolution (reviewed in [24]. When comparing expression between species that differ in color vision, 25.3% (n = 3173) of genes displaying DE (Table 2; SI Table 1.2). For our samples, the investigated hominoid and cercopithecoid species are monomorphic trichromats in all individuals, while the platyrrhine species are polymorphic with trichromatic homozygous females, dichromatic males or dichromatic heterozygous females (Figure 1; SI Table 1.1). However, despite color vision being sexually dimorphic in some of the platyrrhine species, very few genes were significantly differentially expressed in V1 between sexes (0.28% DE, n = 35;

Table 2; SI Table 1.2). There were no neuronal-specific KEGG signaling pathways enriched in genes displaying DE between trichromatic species and polymorphic di/trichromatic species but there were a number of metabolic pathways enriched, including amino acid, nucleotide, glycolysis/gluconeogenesis, and sphingolipid metabolic pathways (SI Table 1.3). Notably, color vision was the only phenotype-DE comparison enriched for the encompassing, large KEGG pathway of "metabolic pathways" and "carbon metabolism" and included 226 genes (SI Table 1.3). This is 8-fold more genes than any other KEGG pathway enrichment, suggesting that there are extensive metabolic differences in the V1 of primates differing in color vision. However, given the clear phylogenetic influence on these samples (Figure 1, SI Table 1.1), we cannot resolve whether these differences are truly due to divergent color visual perception or rather how hominoids and cercopithecoids differ from platyrrhines in V1.

We investigated DE between primarily arboreal and terrestrial species due to the role of V1 in processing signals derived from navigating and perceiving the environment [15]. The primarily arboreal species in our study included all four platyrrhine species and two hominoid species (orangutan and siamang) while the primarily terrestrial species included all four cercopithecoid species and three hominoid species (human, chimpanzee, and gorilla; Figure 1, SI Table 1.1). There were 2,501 genes with DE in V1 associated with differences in primate habitat use (either arboreal or terrestrial; 19.9%, Table 2; SI Table 1.2). Neural KEGG pathways enriched in species differing in habit use included neuroactive ligand receptor interaction, long-term depression, and retrograde endocannabinoid signaling (SI Table 1.3). This phenotype DE comparison was enriched primarily for lipid metabolic pathways (SI Table 1.3). Lipids are important in the brain for long-term energy storage, membrane structure, extracellular signaling, and enhanced propagation of neural signaling (e.g. myelination) [75]. Habitat use DE enrichments for lipid metabolism may suggest that these processes differ significantly in V1 as it responds to visual stimuli from divergent interactions with habitat, such as enhanced V1 processing of depth perception for arboreal versus terrestrial species.

We also explored possible expression differences in V1 that correlate with differences in sociality, using group size as a proxy variable. For the included species, the "small group" included three platyrrhines species (spider monkey, marmoset, and saki monkey), and three hominoid species (siamang, gorilla, and orangutan) with fewer than 20 individuals per group on average and the "large group" included all four cercopithecoid species, one platyrrhine species (squirrel monkey) and two hominoid species (human and chimpanzee) with greater than 20 individuals per group, on average (Figure 1, SI Table 1.1). We found that 13.6% (n = 1704) genes exhibited DE between species that differed in group size (Table 2; SI Table 1.2). Categorical enrichment determined genes displaying DE between species differing in group size had more neural processes enriched than any other DE comparison based upon trait except for diet (see below), including neuroactive ligand receptor interaction, long-term depression, retrograde endocannabinoid signaling, and GABAergic synapse (SI Table 1.3). Of the genes exhibiting DE between large and small group living species, metabolic KEGG enrichments included amino acid sugar and nucleotide sugar metabolism and fatty acid degradation (SI Table 1.3). Because both long-term depression and GABAergic synaptic signaling function in inhibitory signaling [76], enrichments for both of these processes may suggest a significant difference in inhibitory signaling in the V1 among primates differing in group size, which may be due in part to differential visual responses necessary for navigating socially complex environments.

To compare V1 expression between species that differ by diet, we categorized our species as primarily frugivorous, primarily folivorous, or omnivorous [63]. The primarily frugivorous species included two hominoid species (chimpanzee and orangutan) and three platyrrhine species (spider monkey, marmoset, and saki monkey) while we had two primarily folivorous hominoid species (gorilla and siamang; Figure 1, SI Table 1.1). The omnivorous species included all four cercopithecoid species, a single platyrrhine species (squirrel monkey) and a single hominoid species (human; Figure 1, SI Table 1.1). We calculated DE among the

three possible comparisons: folivore-frugivore (4.3% DE, n = 534), folivore – omnivore (5.3%, n = 664), and frugivore – omnivore (7.9%, n = 991; Table 2; SI Table 1.3). Folivorous species differed from omnivorous species in in tryptophan, purine, and steroid metabolic KEGG pathways (SI Table 1.3). Folivorous species differed from frugivorous species in carbohydrate and protein digestion and retinol and tryptophan metabolism (SI Table 1.3). Frugivorous primates displayed much higher numbers of genes exhibiting DE in V1 than frugivore-folivore and folivore-omnivore comparisons, though this may be due to a lack of folivorous species in the available samples (only siamang and gorilla were folivorous; Figure 1; SI Table 1.2). Frugivorous-omnivorous DE was enriched for serotonergic synaptic signaling and was the only diet comparison enriched for any neuron-specific KEGG pathways (SI Table 1.3). Frugivores differed from omnivores metabolically for amino acid metabolism (SI Table 1.3). Given that serotonin levels have an established link to diet [77] and tryptophan is a known precursor for serotonin [78], the finding that both tryptophan and serotonergic signaling are only enriched in DE between species differing in diet may suggest that diet has an impact on V1 serotonin signaling by way of altered tryptophan metabolism.

All trait DE comparisons were enriched for at least one pathway in the metabolic parent categories but not for neural signaling (SI Table 1.3). Our results suggest that neural processes drive differences in V1 gene expression in species that differ in group size, habitat use, and diet while metabolic differences are more responsible for V1 DE between species differing in color vision (SI Table 3). Our approach correlating DE with phenotype suggests that there is a link between expression in V1 and distinctive differences in visually relevant traits.

1.5 Discussion

Although some studies of V1 gene expression compared it both to other regions of the brain, as well as across species [44, 49-51], no other investigations of brain gene expression included nearly the diversity and number of primate species in the current study. As such, this

represents one of the first in-depth investigations of how V1 gene expression differs across a variety of primate species. We found that while global gene expression in primate V1 clusters largely by phylogenetic relatedness, the greatest degree of variation in expression is within hominoids, largely due to divergence of humans and chimpanzees from other one another as well as other hominoid species. Carbohydrate metabolic processes seem to be driving expression differences across the primate tree while neural processes are more conserved in V1. A deviation from this trend is the expression differences between human and chimpanzee. These species are outliers from all other species and display a relatively large amount of intraspecific variation in global gene expression profiles as well as in specific pathways related to synaptic signaling and neuronal cell projections important for maintaining the complex neurological signaling networks key to brain function. In addition to global and interspecific differences in expression, we were able to link V1 DE to differences in visually-relevant phenotypes.

The enrichment for neuron-specific processes in V1 DE, primarily between human and chimpanzee, is consistent with previous findings of enrichments for similar processes, such as synaptic signaling, in differential gene expression that have been reported in various studies of human, chimpanzee, and rhesus macaque brains [39-41]. Also, similar metabolic processes as those enriched here have also been found to differentiate primate brain gene expression in a number of previous studies of a smaller array of primate species [43, 44, 49, 79], including that humans specifically differ largely in metabolic processes related to aerobic glycolysis [40, 41]. Taken together, our results across a broad number of primate species show that these trends for metabolic and neural signaling differences in the brain extend to V1.

Given that V1 is a primary sensory cortex and in view of subtle interspecies differences in visual perception between closely related primates, there might not be much divergence expected over shorter evolutionary times. Consistent with this, we found a significant correlation for greater numbers of genes exhibiting DE between more distantly related species.

Our finding that interspecies DE tends to increase with evolutionary distance has been observed previously in a study of gene expression across ten species (including six hominoids) and six tissue types [45]. Furthermore, we observed a trend for genes involved in metabolic signaling to be differentially expressed consistently across interspecies DE comparisons, regardless of phylogenetic distance, while processes specific to neural signaling were far less common. Because the most metabolically demanding feature of the brain is synaptic transmission, metabolic differences are not mutually exclusive with neural signaling processes, but this is an interesting trend nonetheless. Our findings suggest that altered gene expression of neuron-specific pathways in V1 does not consistently contribute to functional differences between closely-related species, but altered metabolic processes do (e.g. oxidative phosphorylation is enriched in genes exhibiting DE between all cercopithecoid species; SI Table 1.2). The exception to this is that neuron-specific processes were among the most significantly enriched pathways in genes exhibiting DE between human and chimpanzee.

Within primates, selective differences in the genome can be linked to diet and metabolism, suggesting selection has optimized different metabolic processes in lineage-dependent ways [39-41, 80-83]. The human brain is more energetically costly than that of other primates, utilizing ~ 20% of all of the body's metabolic resources, in comparison to non-human primate brains that use less than 10% [84, 85]. Importantly, allometry alone does not explain the increase in human brain appropriation of glucose metabolism at this proportion [86-88]. Our results are consistent with previous findings that humans differ in brain gene expression from chimpanzees for neuron-specific processes related to synaptic transmission and metabolic processes involved in aerobic glycolysis [40, 41, 43, 44, 49, 79]. These data demonstrate that, like other brain regions [38, 41, 47, 48], human lineage-specific neurological changes are present in the visual cortex.

Understanding V1 gene expression differences based upon variation in phenotypes relevant for vision is an area of research that could elucidate the neurological implications of

differences in vision and the selective pressures hypothesized to be linked to these traits. Although there are important caveats to our analysis of phenotype-based DE, primarily an inability to account for phylogenetic influences (see Methods for more details), this still represents a significant effort to link proximate gene expression differences in the brain with evolved variation in ecological traits. DE among species differing in color vision was enriched for a number of metabolic processes but not neural signaling processes, perhaps suggesting that any expression differences in V1 influenced by differences in color visual perception are driven more by metabolic differences than by those of synaptic signaling. However, this phenotype-DE comparison has the strongest influence of phylogeny, and we are not able to parse out if this difference represents that of divergent color perception or a difference between platyrrhines and cercopithecoids and hominoids.

Furthermore, we chose to investigate the link between group size differences and V1 gene expression due to the hypothesized influences of social behavior on primate brain evolution and the possible link to required differences in visual perception among group living primates [25]. While V1 is not explicitly involved in behavioral processes, there were enrichments for pathways known to be involved in behavior. Primates exhibit extensive variation in social traits, and a number of genes (and associated pathways) have been hypothesized to be linked to these behaviors, such as those involved in social bonding or empathy [arginine vasopressin receptor 1A (*AVPR1A*), oxytocin receptor (*OXTR*), and dopamine receptor (*DRD4*)] [89]. Oxytocin signaling changes in response to social interactions is well documented [90, 91] and variations in coding sequence for oxytocin receptor and their associated influence on social behavior have been observed in rhesus macaques [92]. While group size was not enriched for “oxytocin signaling pathway” or “dopaminergic synapse”, the genes *DRD4* and *OXTR* were both differentially expressed in the DE comparison of species differing in group size, alternatively included in the enriched category “neuroactive ligand receptor interaction” (SI Table 3). Our results show that genes putatively important in primate

social evolution and associated processes display significant DE in V1 in primates differing in group size.

Future studies comparing multiple brain regions may determine if the observed trends of altered gene expression between phenotypically distinct primates and between species are V1-specific or a brain region-independent observation. However, the phenotypes investigated here (color vision variation, arboreality and terrestrially, group size and sociality, and diet) do not have implications solely for primate vision evolution and thus are not expected only to impact on visual cortex gene expression evolution and have likely played important roles in other brain regions, and perhaps the brain overall. Additionally, the hypothesized importance of the evolutionary trajectories of these phenotypes are not mutually exclusive of one another - it is most likely that the evolution of trichromatic color vision has been influenced by an interplay of diet, arboreality, social signals, and other important primate traits. Because our methods could not resolve the influences of each of these traits on V1 gene expression evolution, we remained conservative in our interpretations of the link between phenotype evolution and gene expression evolution. Furthermore, given the highly organized and specific cytoarchitecture of the V1 [15], and previously determined influence of structural genes on V1-specifc gene expression in comparison to other regions [49], it is possible that V1-region specific changes in gene expression are linked to maintaining or fine-tuning this cytoarchitecture. However, studies of V1 function are largely limited to tissue-level functional, mechanistic, and cytoarchitectural investigations [93-95] and primarily only in rhesus macaques, resulting in a very limited understanding of the gene expression changes that accompany altered systems-level function. This, in addition to the lack of other brain regions for comparison in our current study across such disparate primate species and the previous findings that V1 gene expression is similar to other brain regions except cerebellum in a limited number of primate species (rhesus macaque, chimpanzee, human) [44], we limit our conclusions about V1 region-specific gene expression and its link to variation in visually relevant phenotypes.

1.6 Conclusion

We investigated the interaction between genotype and phenotype by examining the correlation between gene expression and phenotypic and behavioral traits, including habitat use, color visual system, group size, and diet in a broad sampling of primate species, including many understudied species (e.g. siamang, squirrel monkey, and spider monkey). We determined that neural and metabolic processes known previously to differ between species in other brain regions also demonstrate interspecies and trait-based differences in V1. We show that human and chimpanzee are outliers for V1 gene expression, differing significantly more in neuron-specific processes related to synaptic signaling than other species do. Although these species appear to be the most divergent, they do not exhibit any major differences in the visually-relevant phenotypes investigated here for which we were able to determine significant expression differences. Future studies that include other primate taxa could further investigate the link between differences in primate vision evolution and visual cortex expression differences.

As primates exhibit many unique visual system traits compared to other mammals and understanding the genetic basis for primate visual systems in the V1 region would provide valuable insights into the evolutionary trajectory of these traits. Our data indicate that there is also a correlated difference in gene expression in the initial processing center of visual signals in the brain. We also show that humans differ in brain gene expression in the V1 in a manner similar to other regions. Further investigation of overlap between DE and signals of selection can provide information about which expression changes are adaptive.

1.7 Supplemental Table List

Table Name	Content
Supplemental Table 1.1	Sample-specific information.
Supplemental Table 1.2	Full table of numbers of genes exhibiting DE for all pairwise DE comparisons.
Supplemental Table 1.3	Kyoto Encyclopedia of Genes and Genomes (KEGG) Enrichments for each pairwise DE comparison. All pathways enriched are significant ($p<.05$) and include 5 or greater genes. Specific KEGG term, ID number and parent group are listed.
Supplemental Table 1.4	Gene ontology (GO) cellular component (CC) enrichments for each pairwise DE comparison. All CC's are significant ($p<.05$) and contain 10 or greater genes. Specific CC terms, IDs, and parent group are listed.

1.8 Figures and Tables

Figure 1

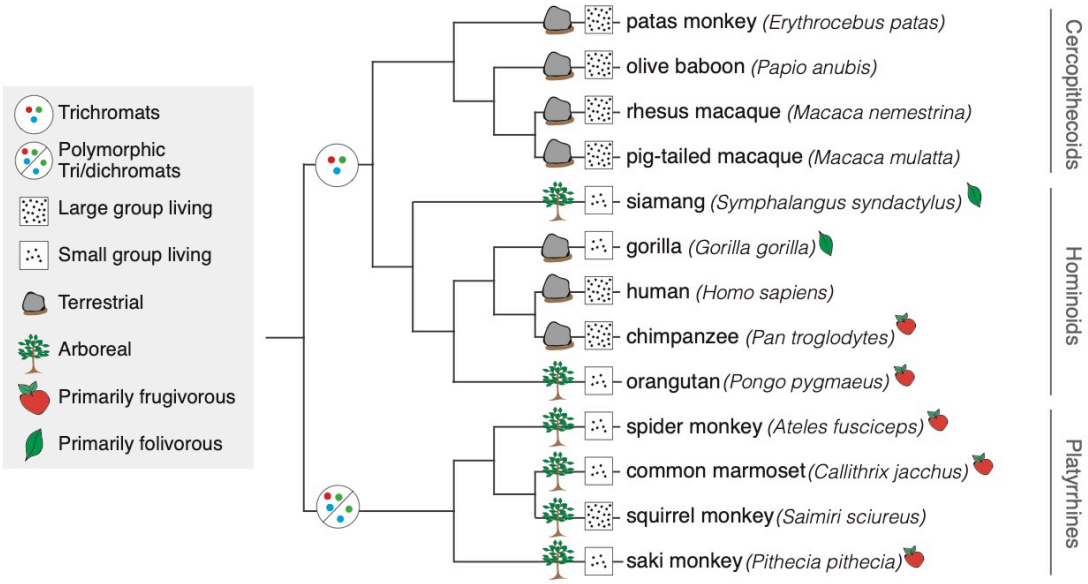


Figure 1 | Phenotypic traits of species for which V1 gene expression was investigated.
Phenotypic traits for color vision, habitat use, and group size are mapped on the phylogeny and diet is depicted to the right of species names. Species without a diet indicator were coded as omnivorous. The tree was generated using 10k Trees Version 3 [59] and Mesquite [96].

Figure 2

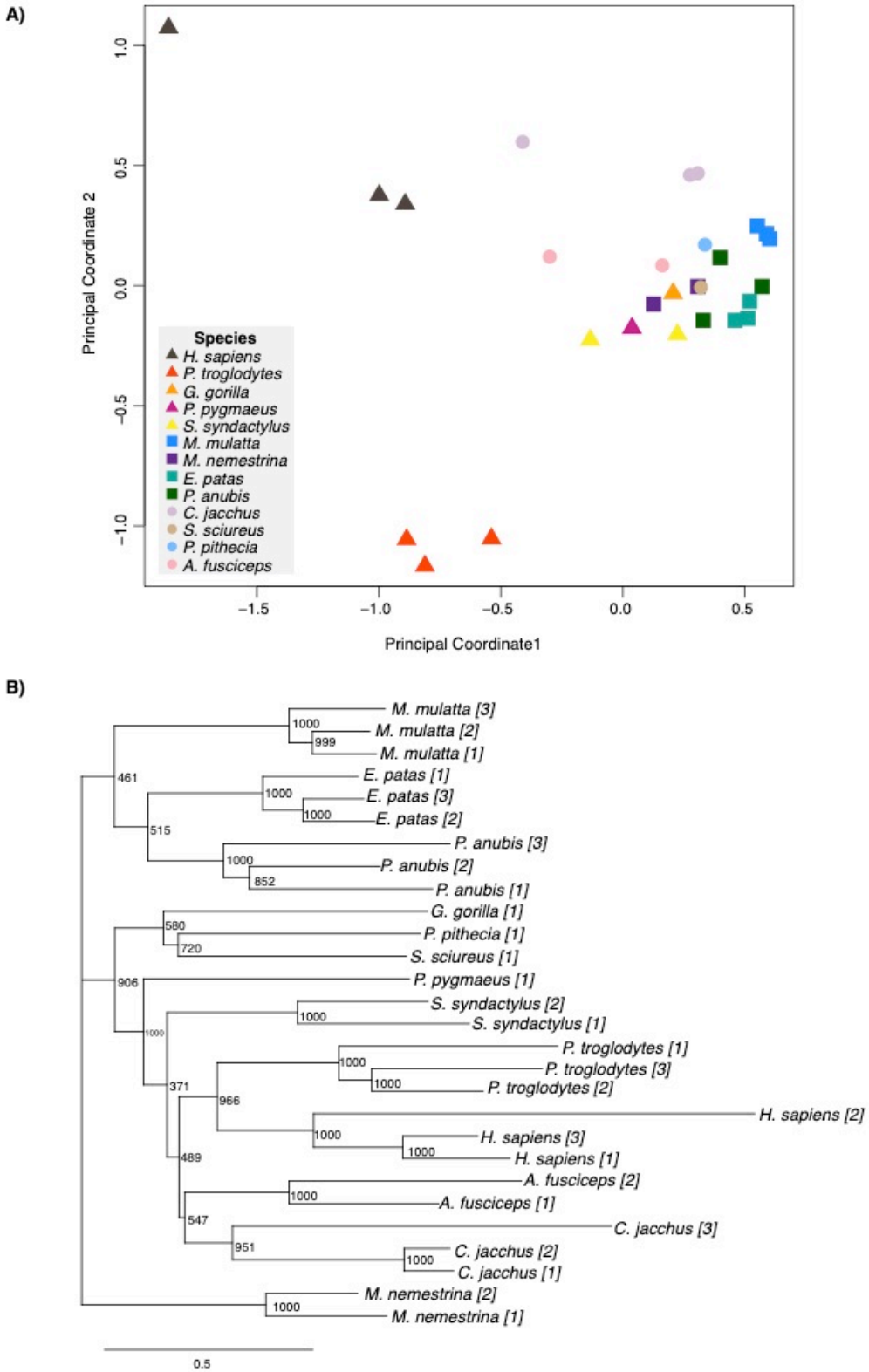

A)
Principal Coordinate 2
Principal Coordinate1
1.0
0.5
0.0
−0.5
−1.0
−1.5
−1.0
−0.5
0.0
0.5
Species
H. sapiens
P. troglodytes
G. gorilla
P. pygmaeus
S. syndactylus
M. mulatta
M. nemestrina
E. patas
P. anubis
C. jacchus
S. sciureus
P. pithecia
A. fusciceps
B)
M. mulatta [3]
M. mulatta [2]
M. mulatta [1]
E. patas [1]
E. patas [3]
E. patas [2]
P. anubis [3]
P. anubis [2]
P. anubis [1]
G. gorilla [1]
P. pithecia [1]
S. sciureus [1]
P. pygmaeus [1]
S. syndactylus [2]
S. syndactylus [1]
P. troglodytes [1]
P. troglodytes [3]
P. troglodytes [2]
H. sapiens [2]
H. sapiens [3]
H. sapiens [1]
A. fusciceps [2]
A. fusciceps [1]
C. jacchus [3]
C. jacchus [2]
C. jacchus [1]
M. nemestrina [2]
M. nemestrina [1]
1000
999
461
515
852
580
720
906
371
966
489
547
951
0.5

Figure 2 | Humans and chimpanzees are the most divergent in V1 gene expression. A) principal coordinate analyses (PCoA) of V1 transcriptomes color-coded by species. Shapes of points indicate clade: triangles for hominoids, squares for cercopithecoids, and circles for platyrrhines. B) Hierarchical clustering of V1 transcriptomes of each sample with bootstrap values and the individual sample number in brackets.

Figure 3

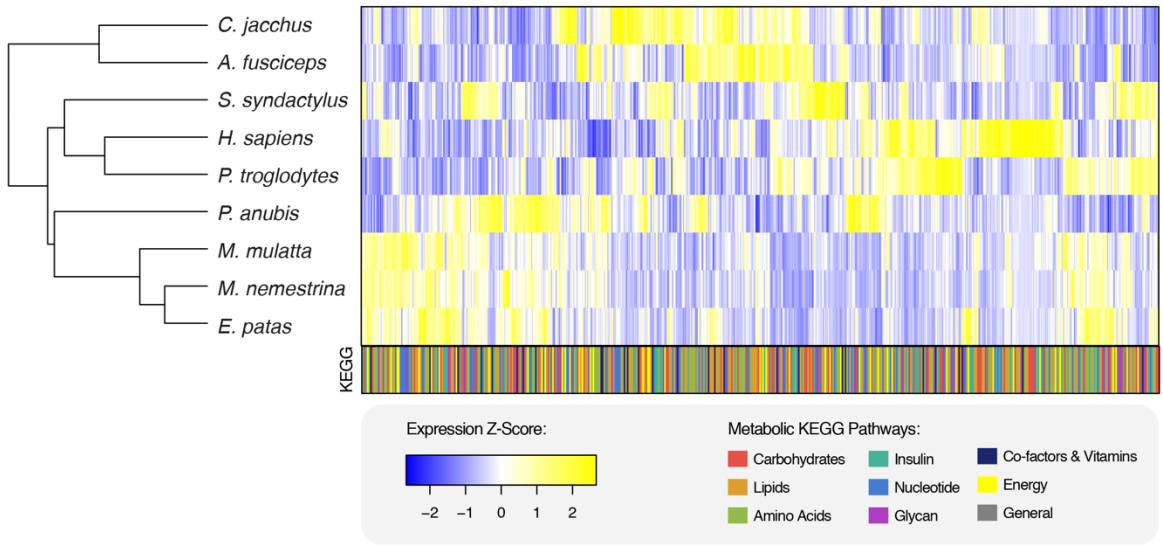

C. jacchus
A. fusciceps
S. syndactylus
H. sapiens
P. troglodytes
P. anubis
M. mulatta
M. nemestrina
E. patas
KEGG
Expression Z-Score:
−2 −1 0 1 2
Metabolic KEGG Pathways:
Carbohydrates
Lipids
Amino Acids
Insulin
Nucleotide
Glycan
Co-factors & Vitamins
Energy
General

Figure 3 | Expression profiles of metabolic genes in primate V1. Clustering of expression profiles of 1,039 metabolic genes in primate V1. Highly correlated genes (columns) cluster together and samples (rows) cluster based on Euclidean distance between expression values. Only species for which there were greater than one sample per species were used. Averages of expression per gene were calculated across replicates per species. The bottom bar represents membership in the color-coded KEGG metabolic pathways for each gene in the heatmap.

Table 1

Interspecies-DE Comparison	KEGG Term	KEGG Parent Category	# Genes	Enrichment P-value	P-Value Rank
Human - Chimpanzee	Cholinergic synapse	Synaptic Signaling	40	0.00001	1
Human - Chimpanzee	GABAergic synapse	Synaptic Signaling	31	0.00011	4
Human - Siamang	Neuroactive ligand receptor interaction	Neuronal Signaling	75	0.00086	3
Human - Siamang	Metabolic pathways	General Metabolism	272	0.00730	9
Human - Baboon	Metabolic pathways	General Metabolism	364	0.00001	1
Human - Baboon	Fatty acid metabolism	Lipid Metabolism	20	0.00558	7
Human - Rhesus macaque	Steroid biosynthesis	Lipid Metabolism	11	0.00118	5
Human - Rhesus macaque	Purine metabolism	Nucleotide Metabolism	55	0.00206	7
Human - Rhesus macaque	Nicotinate and nicotinamide metabolism	Cofactor & Vitamin Metabolism	13	0.00465	10
Human - Marmoset	Metabolic pathways	General Metabolism	349	0.00046	3
Human - Marmoset	Nicotinate and nicotinamide metabolism	Cofactor & Vitamin Metabolism	15	0.00122	6
Human - Spider monkey	Metabolic pathways	General Metabolism	326	0.00024	1
Human - Spider monkey	Steroid biosynthesis	Lipid Metabolism	11	0.00142	3
Human - Spider monkey	Purine metabolism	Nucleotide Metabolism	54	0.00549	7

Table 1 | Metabolic pathways are the most significantly enriched in interspecies DE comparisons with the exception of human-chimpanzee. A table of the top-10 most significantly enriched KEGG pathways in interspecies DE comparisons between human and other species, abbreviated to show only metabolic (green) and neuron-specific (purple) pathways. '# Genes' refers to the number of genes exhibiting DE in that pathway. P-value rank refers to the rank of the enrichment within the top 10 for each comparison. Table is sorted by interspecies DE comparison. Abbreviated from SI Table 3.

Table 2

Phenotype-DE Comparison	# Genes	% Genes
Color vision	3173	25.25
Habitat-use	2501	19.91
Group-size	1704	13.56
Folivore-frugivore	534	4.25
Folivore-omnivore	664	5.28
Frugivore-omnivore	991	7.89
Sex	35	0.28

Table 2 | Proportions of DE genes in primate species V1 vary depending on the compared phenotypes. Pairwise list of phenotype-based DE comparisons. '# Genes' refers to the number of genes exhibiting DE in that pathway. This is a subset of all comparisons made, see SI Table 2 for the full list of genes exhibiting DE for all comparisons. Abbreviated from SI Table 2.

Table 3

Phenotype-DE Comparison	KEGG Term	KEGG Parent Category	# Genes	Enrichment P-value
Metabolic Pathways				
Color vision	Biosynthesis of amino acids	Amino Acid Metabolism	21	0.004
Diet: frugivore -omnivore	Glycine serine and threonine metabolism	Amino Acid Metabolism	5	0.046
Group size	Histidine metabolism	Amino Acid Metabolism	6	0.013
Color vision	Glycolysis Gluconeogenesis	Carbohydrate Metabolism	17	0.030
Diet: frugivore - folivore	Retinol metabolism	Cofactor & Vitamin Metabolism	5	0.030
Color vision	Sphingolipid metabolism	Lipid Metabolism	13	0.028
Diet: folivore - omnivore	Steroid hormone biosynthesis	Lipid Metabolism	6	0.012
Group size	Fatty acid degradation	Lipid Metabolism	8	0.031
Habitat Use	alpha Linolenic acid metabolism	Lipid Metabolism	7	0.030
Habitat Use	Ether lipid metabolism	Lipid Metabolism	11	0.020
Habitat Use	Glycerophospholipid metabolism	Lipid Metabolism	20	0.013
Color vision	Pyrimidine metabolism	Nucleotide Metabolism	26	0.012
Diet: folivore - omnivore	Purine metabolism	Nucleotide Metabolism	12	0.015
Neuron-Specific Pathways				
Diet: frugivore -omnivore	Serotonergic synapse	Synaptic Signaling	12	0.010
Diet: frugivore -omnivore	Synaptic vesicle cycle	Synaptic Signaling	8	0.012
Group size	GABAergic synapse	Synaptic Signaling	14	0.016

Table 3 | Some KEGG pathway enrichments in phenotypic-DE comparisons are not enriched in more than one phenotype-DE comparison. "Uniquely enriched" KEGG terms are shown for each phenotype-DE comparison investigated. KEGG parent categories do overlap, with the exception of "Cofactors & Vitamins". '# Genes' refers to the number of genes exhibiting DE in that pathway. Abbreviated from SI Table 3.

CHAPTER 2

CHIMPANZEE FIBROBLASTS EXHIBIT GREATER ADHERENCE AND MIGRATORY PHENOTYPES THAN HUMAN FIBROBLASTS

2.1 Abstract[2]

Background and objectives: Previous work has identified that gene expression differences in cell adhesion pathways exist between humans and chimpanzees. Here, we used a comparative cell biology approach to assay interspecies differences in cell adhesion phenotypes in order to better understand the basic biological differences between species' epithelial cells that may underly the organism-level differences we see in wound healing and cancer.

Methodology: We used skin fibroblast cell lines from humans and chimpanzees to assay cell adhesion and migration. We then utilized published RNA-Seq data from the same cell lines exposed to a cancer / wound-healing mimic to determine what gene expression changes may be corresponding to altered cellular adhesion dynamics between species.

Results: The functional adhesion and migration assays revealed that chimpanzee fibroblasts adhered sooner and remained adherent for significantly longer and move into a "wound" at faster rate than human fibroblasts. The gene expression data suggest that the enhanced adhesive properties of chimpanzee fibroblasts may be due to chimpanzee fibroblasts exhibiting significantly higher expression of cell and focal adhesion molecule genes than human cells, both during a wound healing assay and at rest.

[2] This chapter is formatted for publication in Evolutionary Medicine and Public Health with a planned submission date sometime in November 2019. The full author list for this manuscript is: Trisha M. Zintel, Delaney Ducey, Alyssa D. Schwartz, Courtney C. Babbitt

Conclusions and implications: Chimpanzee fibroblasts exhibit stronger adhesion and greater cell migration than human fibroblasts. This may be due to divergent gene expression of focal adhesion and cell adhesion molecules, such as integrins, laminins, and cadherins, as well as ECM proteins like collagens. This is one of few studies demonstrating that these divergences in gene expression between closely related species can manifest in fundamental differences in cell biology. Our results provide better insight into species-specific cell biology phenotypes and how they may influence more complex traits, such as cancer metastasis and wound healing.

2.2 Introduction

Although humans and chimpanzees shared a common ancestor relatively recently (~ 5 million years ago), they have largely identical amino acid and protein structure and greater than 98% DNA sequence identity [37, 97]. Despite this high nucleotide conservation, there are many well-documented phenotypic discrepancies related to cognition, behavior, and anatomy, including some intriguing differences that may be biomedically relevant for understanding human-specific disease [97, 98]. Determining specific genetic influences on interspecies disparities in disease etiology is important to understand human-specific diseases [99, 100]. In order to better understand these organism-level differences in complex phenotypes, and their evolutionary basis, we can investigate interspecies differences in basic molecular and cellular biological processes.

One mechanism by which interspecies differences in phenotype may arise from highly similar genomes is through changes in non-coding sequences that influence the expression the same protein-coding genes [37]. Significant changes in gene expression between humans and chimpanzees have been studied previously [38-42, 44-46, 101]; however, studies of the cellular phenotypic impacts of altered gene expression are lacking. One way to start to make those connections is to use available cell lines from multiple species. Fibroblast cells are important for homeostasis and critical for wound healing and often play an integral role in inflammation

during cancer progression [102, 103]. A previous publication found that chimpanzee skin fibroblasts have significantly more focal adhesions per cell than human fibroblasts [104]. Additionally, human and chimpanzee fibroblasts exhibit significant differences in gene expression of adhesion pathways both in fibroblasts grown under normal conditions [101, 105] as well as those exposed to a serum challenge that mimics both wound healing and a cancer response [105]. However, there are no known studies investigating if these differences influence cellular adhesion and migratory phenotypes.

Intercellular signaling facilitated by cell adhesion is vital for many biological processes, such as embryonic development as well as disease states (e.g. metastasis of cancerous cells) and consequently represents a potential mechanism by which functional differences among primates may have evolved. Healing of a wound *in vivo* requires migration of neighboring healthy cells into the wound [106] which is facilitated by disassembly and reassembly of focal adhesion complexes to the leading edge of migrating cells [107-110]. Cell adhesion and migration pathways are important in a number of complex traits, including development and differentiation [110-113] as well as cancer metastasis [114-117]. Focal adhesions (FAs) are large protein complexes that facilitate adhesion of cells to the underlying extracellular matrix (ECM) as well as cell motility by the disassembly and reassembly of these complexes in difference locations along the cell's adherent surface [107-110, 118]. A variety of protein types comprise FAs, including extracellular protein complexes (e.g. zyxin) that interact extracellular matrix proteins (e.g. collagen, laminin) to facilitate adherence as well as transmembrane proteins that function not only in adherence but also in bidirectional signaling across the membrane (e.g. vinculin, integrin) [107].

In order to determine if previous findings of interspecies differences in adhesion complexes and gene expression translate into fundamental differences in cellular phenotypes, we used established cell culture methods to assay adhesion and migration differences between human and chimpanzee primary skin fibroblasts *in vitro*. We hypothesized that there may be

significant differences in cellular adhesion and migration phenotypes between human and chimpanzee fibroblasts that correlate with the previously observed differences in focal adhesion gene expression. We investigated adhesive properties using two different cell adhesion assays, that measured timing and strength of fibroblast adherence and migration phenotypes by comparing migration patterns, speed, and distance of cells over time in two conditions. We also used a classic scratch test that mimics a wound and the subsequent healing process *in vitro* [119, 120]. Finally, in order to examine underlying gene expression differences in adhesion and migration signaling pathways between species at the cellular level, we incorporated previously published data on differential gene expression between human and chimpanzee fibroblasts during an assay that mimics both wound healing and cancer progression [105].

2.3 Methodology

2.3.1 Samples

Primary fibroblasts from two individuals per species were obtained from Coriell Biorepository (Camden, NJ) (human lines AG10803 and AG09605, chimpanzee lines S006007 and S008956). Cells were cultured with minimum essential media (MEM) supplemented with 10% fetal bovine serum (FBS), 1% L-glutamine, & .1% antibiotic at 37C and 5% CO_2, according to Coriell protocols. Cells were sex-, age-, and body-region- matched.

2.3.2 Cell adhesion assay

Fibroblasts at 80-90% confluency in 6-well plates were dyed with a non-toxic cytoplasmic dye for ten minutes at a concentration of 1 μM (CellTracker, catalog #C7025, Invitrogen). A live cell nuclear stain was then added for twenty minutes, according to manufacturer's instructions (NucBlue, #R37605, Invitrogen). Fibroblasts were subsequently passaged using .5% trypsin and seeded as evenly as possible into eight individual wells of a 12-well plate. At four timepoints (.5, 1, 1.5, 2 hours), wells were rinsed to remove non-adherent cells and fluorescent images were taken. To maximize accurate counts of cells/well, each well

was divided into two quadrants, and for each timepoint, one photo per quadrant was taken. Cell counts were obtained using FIJI Is Just ImageJ (FIJI) [121] to quantify the fluorescently dyed cells. We calculated a mean of the two counts per timepoint. Experiments were conducted simultaneously in triplicate for all four cell lines. DNA concentration was used to control for differences in cell number by dividing cell counts at each time point by the DNA concentration an equal aliquot from time zero.

2.3.3 Flow challenge to assay cell adhesion and migration

Fibroblasts of each species were assayed for differences in cell adherence and migration using a microfluidic device equipped with a time-capture, environment-controlled microscope system, developed in Shelly Peyton's laboratory (University of Massachusetts Amherst). Fibroblasts at 80% confluency were trypsinized and 1,000-150,000 cells were seeded into the flow cell device to monitor cell adhesion and migration over time (referred to further as adhesion-in-flow experiments). The flow cell device uses two reciprocating syringe pumps connected via tubing to a narrow chamber over which cells suspended in media were pumped into at a rate of 50 μL/minute for 6 hours. Photographs of the fibroblasts adhering and migrating in the flow cell were taken every 15 minutes over the 6-hour time course. Individual fibroblasts were identified and assessed over the 6-hour observation period and were traced using FIJI consecutively across this time period to obtain measurements of initial fibroblast adherent surface area and subsequent cell spread via centroid X and Y coordinates over time [121]. These data were used to plot migration paths of individual fibroblasts per species (SI Figure 1A). Distance traveled of individual fibroblasts was calculated with a standard distance equation ($D = [(x_2-x_1)^2 + (y_2-y_1)^2]$) [122]. Speed was calculated in microns/minute using total distance traveled divided by time spent adherent for each individual cell [123]. Adherence duration was calculated by averaging total lengths of time individual fibroblasts spent adherent. Timing of fibroblast adherence was determined by comparing the proportion of fibroblasts adherent at a particular time point (i.e. .5 hours, 1 hour) to that of the total number of fibroblasts that adhered

over the entire time course. However, due to significant cell death, sufficient cell numbers (n = 25-30, per experiment) were obtained only for a single experiment per cell line (two individuals per species) and we could not obtain the planned technical replicates per cell line we had attempted. However, per species, we analyzed a minimum of 80 cells. Significant cell death likely occurred due to the required changes from typical mammalian cell culture needed for this assay, which included a transition from CO_2-independent media to CO_2-dependent media in addition to consistent sheer force on these primary, un-immortalized cells due to the reciprocating flow across the surface for the duration of the experiment.

2.3.4 Wound healing assay:

In order to assess interspecies differences in cell migration, we conducted a wound healing assay as previously described [119, 120]. Briefly, 70-90% confluent fibroblasts were dyed with a non-toxic cytoplasmic dye as previously described and scratched ("wounded") using a p200 pipette tip. Each well was then rinsed and media was replaced to remove cell debris prior to taking fluorescent images (time 0). The same region of the well was subsequently imaged at 24, 48, and 72 hours. We used CellProfiler (cellprofiler.org) to quantify the number of cells that migrated into the wound and the area covered by cells over time [124]. Experiments were conducted in triplicate for all four cell lines.

2.3.5 Identifying candidates for altered cell adhesion and migration from a study of differential expression of genes during a wound healing response

In order to determine gene expression differences that may correspond to the interspecies cellular differences in adhesion in our experiments, we used data produced by Pizzollo et al. [105], a published study from our laboratory, that determined genes exhibiting significant different expression (DE) (q-value < 0.05) between human and chimpanzee fibroblasts undergoing a wound healing response. This study investigated global open chromatin and transcription (via DNAse-Seq and RNA-Seq, respectively) in the same cell lines investigated here. Pizzollo et al. [105] found genes exhibiting significant DE between species'

fibroblasts at rest (unchallenged) and those undergoing a similar wound-healing response to the one used in our experiments. For our analysis, we obtained these lists of significantly differentially expressed (DE) genes from [105] and narrowed them down to those that are members of the Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways [64] for "focal adhesion" and "cell adhesion molecules" (SI Table 2.1). We then compared whether they were DE between species fibroblasts at rest, during the wound healing response, or both to determine putative genetic differences underling the cellular phenotype differences in adhesion and migration observed in our experiments.

2.4 Results

2.4.1 There are phenotypic differences in cell adhesion between human and chimpanzee fibroblasts

In order to investigate if cellular phenotypes related to adhesion differ in manner consistent with previously findings of significant differences in focal adhesions between human and chimpanzee fibroblasts [104], we wanted to determine if there were differences in timing of fibroblast adherence to a surface. In order to determine if there was a difference between species in incidence of adhesion over time, we conducted multiple assays of fibroblast adhesion and compared across species. Initially, we conducted a straightforward investigation of cell adherence by seeding trypsonized fibroblasts into multiple wells and then removing excess, suspended cells at sequential timepoints (.5, 1, 1.5, and 2 hours post-seeding) to count the number of cells that had already adhered (Figure 4A). We expected that if chimpanzee cells adhered more strongly than human cells, we would see more chimpanzee cells adhere during earlier timepoints. We observed greater numbers of chimpanzee cells adhering than human fibroblasts at all timepoints (Figure 4A). There was a statistically significant difference between species (two-way ANOVA $F=13.59$, $p = 0.0142$) but not time (two-way ANOVA $F=.62$, $p = 0.4665$) (Figure 4A).

We next used a more complex assay for cell adhesion using a flow cell device that continuously flowed media with suspended fibroblasts back and forth over a plastic surface, allowing us to monitor adhesion via photographs over a six hour time period (further referred to as the 'adhesion-in-flow' assay). One benefit to this approach was that it allowed us to better control cell number across experiments (100,000 per experiment for human cells, 150,000 for chimpanzee cells). Individual fibroblasts were identified in these sets of images and traced over time to obtain counts of adherent cells. We calculated percentages of observed adherent cells at each time point as well as the length of time cells of each species remained adherent. When we calculated the mean total length of time fibroblasts were adherent over the six hour time course, chimpanzee fibroblasts were adherent for significantly longer (5.4 hours) than human fibroblasts (4.56 hours, $p < .001$; Welch 2-sample T-test; Figure 4B) and adhered in greater numbers within the first hour than human cells (Figure 4C). These results demonstrate chimpanzee fibroblasts adhering faster than human fibroblasts and may mean chimpanzee cells exhibit firmer cell adhesion than human fibroblasts.

2.4.2 Chimpanzee fibroblasts do not appear to inherently migrate faster or farther than human fibroblasts but do exhibit faster *in vitro* wound healing

The increased number of focal adhesions in chimpanzee fibroblasts may also confer differences in cell motility due to dynamic nature of disassembly and reassembly of focal adhesion complexes during cell migration [107-110]. We investigated if there were inherent differences in cell migration between species. Individual fibroblasts from the adhesion-in-flow experiments were traced using FIJI software throughout the full six-hour experiment to obtain area measurements and X & Y coordinates. In order to determine if there were differences in migration patterns between fibroblasts, we used the centroid X and Y coordinates obtained from the six-hour set of photographs to construct a plot of migration patterns between species (SI Figure 6A). There appears to be little interspecies difference in migration patterns (SI Figure 6A), though, qualitatively, there does seem to be more movement along the Y-axis for human

cells than chimpanzee cells. Using these coordinates, we calculated distance traveled and speed of migration for each individual fibroblast over time. There was not a significant difference between species in mean distance migrated (SI Figure 6B) or speed of migration (microns/hour, SI Figure 6C). There was also no difference in cell spread (area/time, data not shown).

Given this fundamental biological link between focal adhesions and cell motility as well as the evidence for greater adhesive properties of chimpanzee cells, we further assessed cell migration by conducting wound healing experiments, which allow us to test a more biologically relevant scenario of cell migration (SI Figure 6A). Wound healing assays, also known as 'scratch tests' are a straightforward way to investigate cell migration *in vitro*, and involve making a scratch in a monolayer of cultured cells and monitoring the scratch over time as cells respond to the gap [119, 120] (Figure 5A). We found a statistically significantly increase in number of chimpanzee fibroblasts that migrated into the wound (two-way ANOVA $F=30.36$, $p = 1.66e-06$; Figure 5B). There was also a statistically significant difference in cell count for both species in original scratch over time (two-way ANOVA $F=62.02$, $p = 5.23e-10$; Figure 5B). This resulted in chimpanzee fibroblasts covering a greater percentage of the initial wounded area than human fibroblasts during the first 24 hours post-wounding (Figure 5C). However, while there was a statistically significant difference in area covered over time (two-way ANOVA $F=14.20$, $p = 0.000475$), there was not between species (two-way ANOVA $F=1.18$, $p = 0.283221$). These results demonstrate that larger numbers of chimpanzee fibroblasts migrate into an *in vitro* wound than human fibroblasts do.

2.4.3 Human and chimpanzee fibroblasts increase expression of distinct FA and CAM genes normally and during wound healing

In order to see how gene expression diverges during the wound healing response between species, we harnessed previously published data from our lab that compared gene expression profiles via RNA-Seq of human and chimpanzee fibroblasts during a serum challenge that mimics both wound healing and metastasis, using many of the same cell lines as

those investigated here [105]. In this study, while human and chimpanzee fibroblasts both exhibited gene expression profiles characteristic of wound healing throughout the serum challenge, there were a number of genes that were significantly differentially expressed (DE) between species in the control, unchallenged samples (pre-challenge) as well as during the wound healing response (SI Table 2.1) [105]. In order to determine what interspecies differential expression might explain differences in adhesive properties between normal primary fibroblasts, and may be coinciding with our observed increase in cell migration during the wound healing assay, we identified genes involved in the KEGG "focal adhesion" (FA) and "cell adhesion molecules (CAMs)" pathways from the lists of genes Pizzollo et al. (2018) identified as significantly differentially expressed (DE) between species' in the normal, unchallenged fibroblasts and during the serum response (12 and 24 hours post-serum reintroduction) [105]. Cell adhesion and migration is modulated by interplay between cell adhesion molecules and focal adhesions [125], and understanding the changes in their expression during a wound healing response can provide insight into the genetic influences on these cellular level phenotypes. There were 27 FA and 31 CAM genes DE between species at any timepoint (SI Table 1). We focused on those that belonged to functionally distinct adhesion related gene families in order to determine if there were interspecies differences expression in these gene families in unchallenged or serum-challenged fibroblasts. Several cadherin genes were DE between species at all timepoints, with CDH2 and CDH4 more highly expressed in humans and CDH1 and CDH3 more highly expressed in chimpanzees (SI Table 2.1,2.2). Chimpanzee fibroblasts exhibited significantly greater expression of the collagen gene COL4A6 at all time points and COL9A3 during the serum challenge while human fibroblasts only significantly increased expression of the collagen genes COL6A1 and COL6A3 during the serum challenge (SI Table 1,2). Chimpanzee fibroblasts had significantly higher expression of integrins at all time points (ITGA7 and ITGA9, SI Table 2.1,2.2) while human fibroblasts did not. Human fibroblasts increased LAMA4 expression during the serum challenge while

chimpanzee fibroblasts always exhibited higher expression of LAMC3 and increased LAMA3 and LAMC2 when unchallenged (SI Table 2.1,2.2). Chimpanzees exhibited greater expression of the syndecan gene SDC2 at all time points, but unchallenged human cells upregulated SDC3 and while unchallenged chimpanzee cells upregulated SDC4 (SI Table 2.1,2.2). Interestingly, chimpanzees exhibited higher expression of a number of these genes at all timepoints, regardless of challenge-state (Table 4). This highlights that humans and chimpanzees are upregulating distinct genes from of FA/CAM protein families that may contribute to interspecies differences in adhesion and migration cellular phenotypes.

2.5 Discussion

Our results demonstrate that chimpanzee fibroblasts adhere faster using multiple cell adhesion assays (*in static* and *in flow*) (Figures 4 & 5). We consistently observed the most stark interspecies differences at earlier time points: within the first hour for *in static* experiments (Figure 4A), the first two hours (in terms of cell count) during the adhesion-in-flow experiments (Figure 4C) likely contributing to the longer overall time adherent for chimpanzee cells (Figure 4B), as well as in the first 24 hours of the wound healing assay (Figure 5B, C). Significance in interspecies difference in cell count dropped at later time points: post-3 hours for adhesion-in-flow experiments (Figure 4C) or beyond 24 hours in the wound healing assay (Figure 5C) or became unreliable to assess in *in static* adhesion assay (Figure 4A, decrease in cell number at later times). These results in light of previous work demonstrating greater numbers of focal adhesions in chimpanzee fibroblasts [104] may suggest that biological processes dependent on rapid adherence of cells surface may differ significantly between human and chimpanzees, but those relying on more gradual adherence or cell migration may ultimately end up with similar phenotypes across species.

Given that stronger cell adhesion is associated slower cell migration [125], it is at first potentially paradoxical that we appear to observe stronger adhesion in chimpanzee fibroblasts

during the early timepoints of our adhesion experiments (Figure 4), no difference in overall cell speed or distance between species fibroblasts (SI Figure 1) but apparently faster cell migration of chimpanzee fibroblasts in our wound healing assay (Figure 5B). The biological context of these different assays likely plays a role. Our *adhesion-in-flow* experiments were stressful for these primary cells regardless of species, exemplified by high cell death (see Methods). These also assess cell migration in different contexts, one with initial cell adherence followed by migration in a challenging, flowing environment (*adhesion-in-flow*) while the wound healing scratch test assesses migration of already adherent cells into newly unoccupied space. Given that scratch/wound healing assays are a commonly used, classic assessment of cell migration, we feel these results are robust.

It appears that human fibroblasts may be larger than chimpanzee fibroblasts during our *in vitro* wound healing assay (Figure 5A). This may be responsible for the lack of significant difference between species in area covered by species in a two-way ANOVA over all times (Figure 5C). However, there is significant difference in percent of area covered at 24 hours post-scratch (Figure 5C). Our *in vitro* assessment of wound healing is not sufficient to ultimately determine if interspecies differences in cell size contribute more strongly to the anecdotal evidence of faster wound healing than number of cells migrating into wound. However, the previous work that determined that chimpanzee fibroblasts have greater numbers of focal adhesions per cell than humans controlled for differences in cell size [104], and so this does not ultimately negate these results. However, though it is not within the scope of the current study, an interesting future direction could be if the there is a larger trend toward larger cells in humans and if so, if there is any corresponding biological advantage in processes such as wound healing or cell motility.

Harnessing RNA-Seq data of the same cells lines during a similar wound-healing and cancer-like *in vitro* assay [105] allowed us to examine gene expression changes coinciding with interspecies differences in adhesion and migration during a wound healing response. We were

unable to use these data to compare gene expression differences in vinculin to corroborate the finding by Advani et al. [104] of interspecies differences in vinculin staining because it was not one of the genes included in Pizzollo et al. [105]. However, it seems that interspecies differences in adhesion and migration may largely be driven by the constitutively higher expression of other CAM and FA genes in chimpanzee fibroblasts, at rest and during a wound healing response (Table 4). Cell adhesion and migration is modulated by interplay between cell adhesion molecules and focal adhesions [125] and concerted increases in expression of both likely is responsible for important interspecies differences in adhesion and cell motility. Integrins are a family of heterodimeric proteins transmembrane proteins with a critical roles in focal adhesion complexes as bidirectional signaling between the interior of the cell and the ECM, and vice versa [107, 108, 126]. It is interesting that chimpanzee fibroblasts but not human fibroblasts exhibited greater expression of integrin proteins regardless of serum status (Table 4, SI Table 2) [105]. Given that integrins are directly involved in communicating signals bidirectionally across the plasma membrane, this may indicate that chimpanzee fibroblasts have enhanced FA signaling compared to human fibroblasts, which may contribute to the phenotypes observed here of enhanced adhesive and motile properties in chimpanzee fibroblasts.

One of several of the major disparities in disease occurrence between humans and chimpanzees is that of incidence and progression of metastatic cancer. Though we clearly lack the same extensive population-level understanding of cancer incidence in chimpanzees, the data we do have from necropsies of captive chimpanzees indicates that epithelial cancer occurs at a rate of $\leq 4\%$, in sharp contrast to the $> 20\%$ of human deaths due to epithelial neoplasms, such as breast or lung carcinomas [100, 127-130]. Of further biomedical relevance, there is anecdotal evidence that humans and chimpanzees also differ in wound healing [131]. Intriguingly, these two disease states are linked – tumors have been described as wounds that do not heal [102] and gene expression profiles of cancerous tissue mimic that of tissue undergoing wound healing [132, 133]. The interspecies disparities in wound healing and cancer incidence

may be due to alterations in the basic biological processes common to both of them, namely, cell adhesion and migration. Cancer is complex disease that results from aberrations of some of the most fundamental biological processes, including cell growth, proliferation, and migration. Fundamental differences in these processes between humans and chimpanzees may help explain to some extent the significant differences in cancer-related deaths. Several integrins expressed in epithelial cells influence metastasis, though whether increased or decreased expression in cancerous compared to normal tissue is inconsistent across integrins and cancer types [134]. Both of the integrins with increased expression in chimpanzee fibroblasts have been linked to metastasis in cancerous cell lines [135-138]. Reduced expression of ITGA7 was implicated with greater cell motility and metastasis [139] and thus there is potential that increased integrin expression in chimpanzee fibroblasts does contribute to decreased cell motility and potentially, decreased metastasis of cancerous cells. All of the cadherins up regulated in human (CDH2, CDH4) and chimpanzee (CDH1, CDH3) are implicated in cancers of various kinds (Roy 2014). CDH1 and CDH3 expression typically overlap (Roy 2014).

Furthermore, extracellular matrix proteins important for cell adhesion and migration include laminins and collagens. Collagen IV is the main collagen composing the basement membrane. COL4A6 was upregulated in chimpanzees at all timepoints (SI Table 2.2). However, during the serum challenge, humans upregulated COL6A1, COL6A3, and LAMA4 while chimpanzees upregulated COL9A3 in addition to their always upregulated COL4A6. This suggests that during a wound healing response, both species will upregulate secretion of basement membrane proteins of the ECM of similar families, but not the same genes. Finally, syndecans are transmembrane glycoproteins whose extracellular domains often act as co-receptors [140]. Interestingly, in normal fibroblasts, humans upregulated syndecan-3 (SDC3), a syndecan primarily found in adults in neuronal tissues (Table 4, SI Table 2.2) [140]. Chimpanzees upregulated SDC2 at all timepoints and SDC4 in normal fibroblasts, both of which are broadly expressed syndecans across multiple tissue and cell types [140]. This is in

line with previous findings that altered gene expression in humans is enriched for neural processes [140]. This may represent a shared evolutionary pressure for increased neural activity.

2.6 Conclusions and Implications

The degree of change in coding sequence and protein function between humans and chimpanzees has long known to be insufficient to explain the stark phenotypic differences between these species [37, 99, 100]. Alterations in gene expression are known to influence human-specific phenotypes, including some that influence disease [97-99, 141]. However, this is one of few studies demonstrating that these divergences in expression can manifest in fundamental differences in cell biology. We demonstrate that there are functionally important differences at the cellular level between species corresponding that correspond with significant interspecies divergence in gene expression in wild-type unchallenged fibroblasts as well as those undergoing an *in vitro* mimic of both wound healing and cancer. Our study represents one of relatively few investigations into the basic cellular differences between species that may allow us to better understand evolutionary of complex disease traits

2.7 Figures and Tables

Figure 4

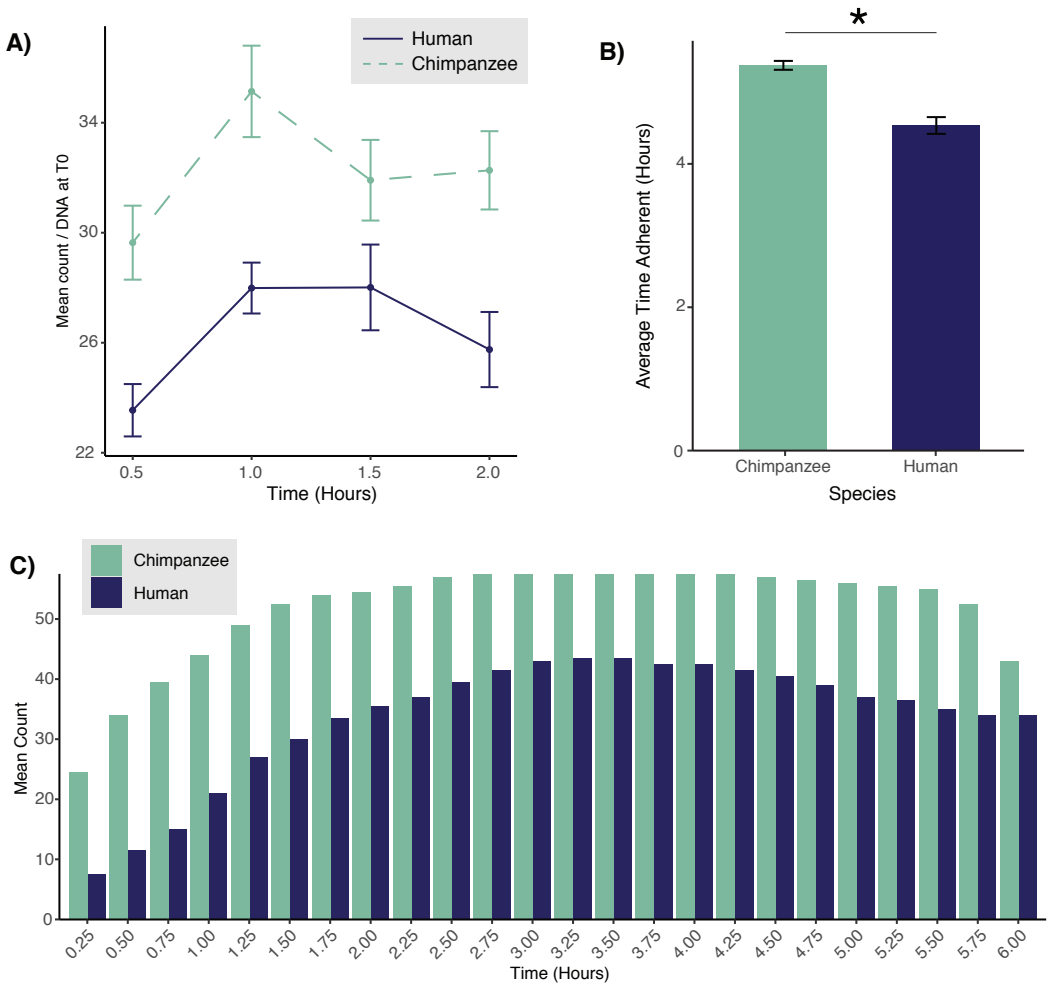


Figure 4 | Adhesion differences between human and chimpanzee fibroblasts. Fibroblasts from two individuals per species were assayed for adherence *in static* (A) and *in flow* (B, C). A) Adhesion-in-static assay. Mean count per timepoint normalized for differences in cell number by dividing raw cell counts at each time point by the DNA concentration from time zero. B) Comparison of the duration of fibroblast adherence *in flow* (per species) ($p < .001$; Welch 2-sample T-test). C) Comparison between species of the timing of fibroblast adherence over the 6-hour time course.

Figure 5

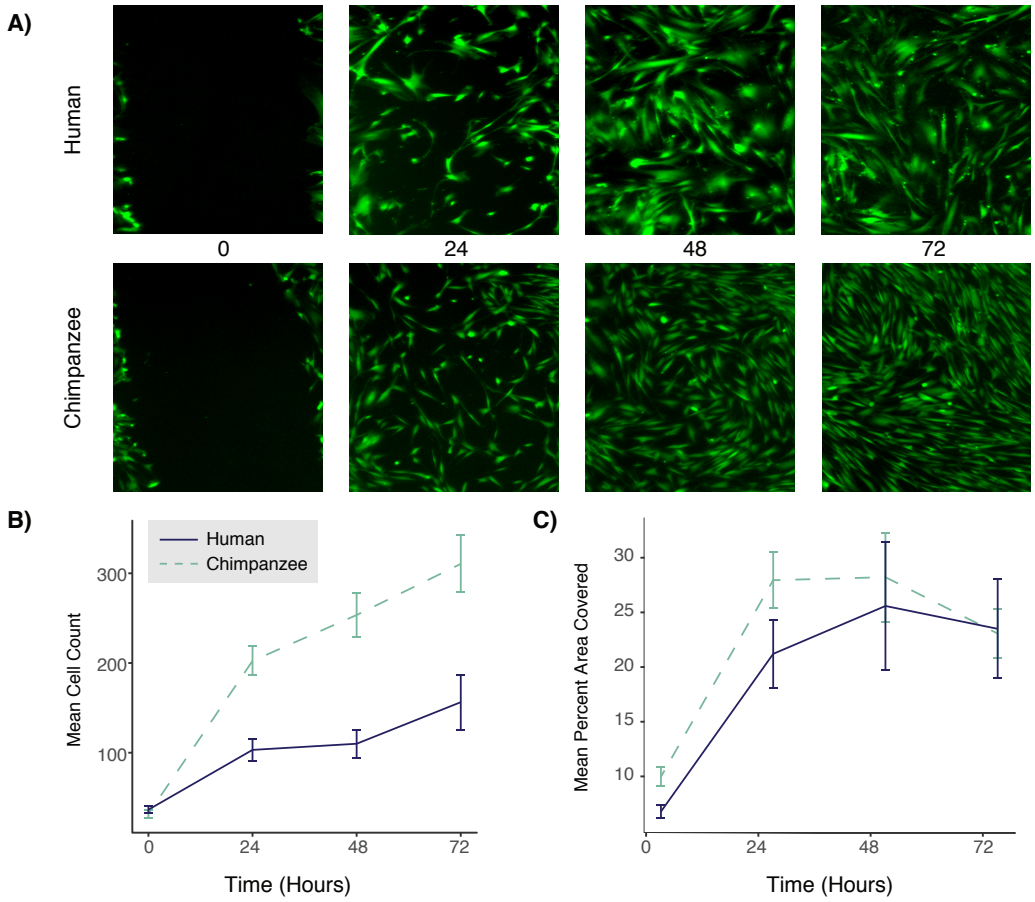

A)
Human
Chimpanzee
0
24
48
72
B)
Human
Chimpanzee
Mean Cell Count
300
200
100
0
24
48
72
Time (Hours)
C)
Mean Percent Area Covered
30
25
20
15
10
0
24
48
72
Time (Hours)

Figure 5 | Interspecies differences in wound healing. A) Representative images of human and chimpanzee fibroblasts undergoing a wound healing response after initial scratch (0), and 24, 48, and 72 hours post-scratch. CellProfiler was used to quantify the number of cells that moved in to fill in the scratch (B) and the change in area covered by cells (C) at 0, 24, 48, and 72 hours post-scratch.

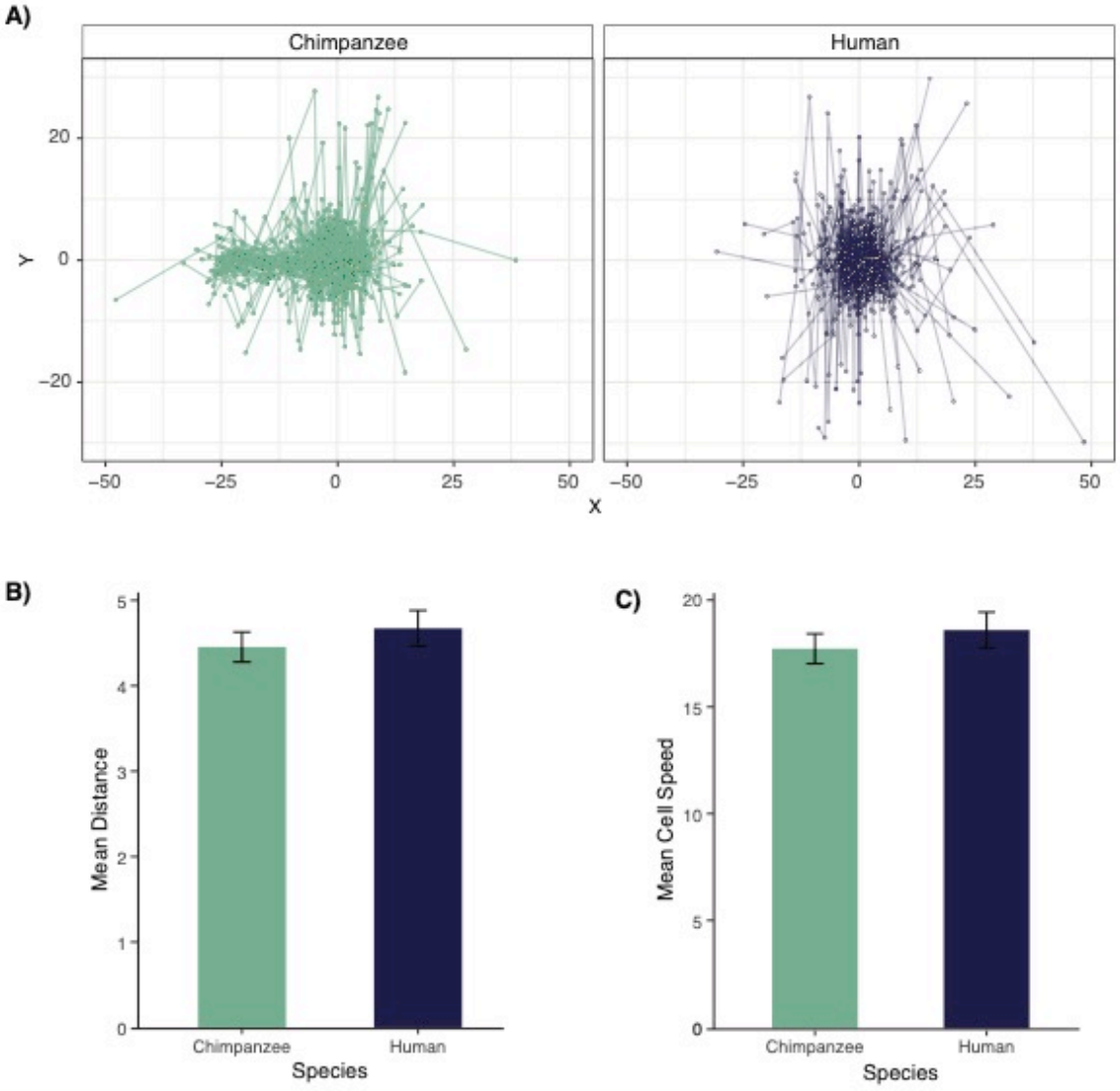
Table 4

DE Time	Higher in Human	Higher in Chimpanzee
During serum challenge only	COL6A1**, COL6A3**, LAM4**	COL9A3**
At all time points (before & during)	CDH2*, CDH4*	CDH2*, CDH3*, COL4A6**, COL9A3**, ITGA7***, ITGA9***, LAMC3**, SDC28

Table 4 | FA and CAM genes DE between species during a serum challenge**.** Comparison of DE patterns of major adhesion gene family genes during serum challenge or at all time points. Asterisks indicate membership in KEGG pathway: * Cell Adhesion Molecules (CAM), ** Focal Adhesion (FA), *** both.

2.8 Supplemental Figures

SI Figure 6



SI Figure 6 | Human and chimpanzee fibroblast migration. Migration patterns of all observed human and chimpanzee fibroblasts plotted using X & Y coordinates obtained from FIJI (A). Species means for total distance traveled (B) and speed (C) of individual fibroblasts.

CHAPTER 3

INVESTIGATING CELL-TYPE SPECIFIC SHIFTS IN METABOLIC GENE EXPRESSION BETWEEN HUMANS AND CHIMPANZEES

3.1 Abstract[3]

The human brain is more energetically costly than that of other primates, utilizing ~ 20% of all of the body's metabolic resources, while chimpanzee brains use less than 10%. Previous work on whole brain tissue has consistently determined significant differences in metabolism between species; yet, a cell-type specific comparison is necessary to understand distinct cellular contributions to interspecific differences in neurological function. Here, we conducted a cell-type specific investigation of neural differences between humans and chimpanzee by conducting RNA-Seq on neural progenitor cells (NPCs), neurons, and astrocytes generated from induced pluripotent stem cells (iPSCs) from both species. Differential expression (DE) analyses demonstrate that the greatest proportion of genes exhibiting DE are between species' astrocytes (12.2% of all expressed genes), followed by NPCs (8.57%), and neurons (5.8%). Categorial enrichment analyses of genes exhibiting DE show that all three neural cell types display interspecies gene expression enriched for metabolic processes, driven primarily by higher expression of metabolic genes in human cells. Gene set enrichment analyses reveal that human neurons and astrocytes are enriched for greater glucose transport while chimpanzee neurons and astrocytes exhibit greater lactate transmembrane transport. We also determined cell-type by species differences in gene expression for nuclear- and

[3] The full author list for this publication: Trisha M. Zintel, Jason Pizzollo, Courtney C. Babbitt

mitochondrially-encoded subunits of the protein complexes important for oxidative phosphorylation. This work circumvents the technical and ethical hurdles of investigating rare cell types in primates to provide exciting insight into the cell type specific metabolic changes that were necessary to support evolution of the human brain.

3.2 Introduction

Though primates exhibit widespread variation in many phenotypes, including anatomy, behavior, and cognition, the extent of these phenotypic differences is not substantially larger than differences in genome sequence [97, 98]. One of those traits that defines primates is a significantly larger brain relative to body size, for which humans exhibit the greatest amount of difference. In 1975, decades prior to sequencing the entire human genome, King and Wilson hypothesized that phenotypic differences amongst humans and chimpanzees must largely be due to changes in gene expression, as coding sequence differences were not nearly divergent enough to account for those differences [37]. Subsequent investigations into differential gene expression across species, and primates in particular, have reinforced the hypothesis that non-coding, *cis*-regulatory changes are critical for these large differences in phenotype [40-42]. Recent work has determined that changes in gene expression in the brain have been found to be both functionally and evolutionarily important [40, 41, 44, 51, 83, 142]. Changes in non-coding regions of the genome identified to be under positive selection in humans are enriched for biological processes and molecular functions that are neural-specific (e.g. synaptic signaling) and those related to metabolism (e.g. carbohydrate and lipid metabolism) in various studies of human, chimpanzee, and rhesus macaque brains [39-41].

Within primates, selective differences in the genome can be linked to diet and metabolism, suggesting selection has optimized different metabolic processes in lineage-dependent ways [39-41, 80-83]**.** The human brain is more energetically costly than that of other primates, utilizing ~ 20% of all of the body's metabolic resources, in comparison to non-human

primate brains that use less than 10% [84, 85]. Importantly, allometry alone does not explain the increase in human brain appropriation of glucose metabolism at this proportion [86-88]. There is evidence that sheer increase in neuron number can explain at least part of the energetic demand of the human brain [143]. However, interspecies differences in the contribution of metabolism of astrocytes versus neurons to metabolic capacity at the organ level remain largely un-explored.

Many of these changes in brain metabolism have been hypothesized to coincide with other trait changes, particularly those related to shifts in diet known to be important in hominin evolution, such as an increase in meat products, increased quality of food, and agriculture [83, 144-149]. A recent study investigating an extensive number of primates is consistent with the idea that diet is a better predictor of brain size than social group structure [63]. The expensive-tissue hypothesis posits that a trade-off in energy allocation for the development of a larger, metabolically demanding brain in primates coincided with a reduction in gut tissue [144, 150-152]. Similarly, an increase in energy-storing tissue (adipose tissue) at the expense of energetically demanding muscle tissue may have also allowed for greater allocation to a larger brain [153-155]. There is evidence that the higher metabolic costs of the human brain influences the protracted development of body growth rate [142]. Evolutionary differences in brain metabolism are a subset of studied differences in metabolic traits that exhibit intriguing differences across primate species. Primates exhibit a lower total energy expenditure (TEE) to body size ratio than non-primates, and furthermore, humans have greater TEE than closely related great ape species (chimpanzees, bonobos, gorillas, and orangutans) due in large part to increased basal metabolic rate (BMR), all of which is consistent with the reallocation of metabolic investment to brain [151, 156]. These *in vivo* (whole organism) studies further suggest an important link between evolutionary differences in metabolism and the uniqueness of the primate brain.

Similar to organism-level investigations, there is also molecular evidence supporting the evolution of metabolic processes (e.g. oxidative phosphorylation) in the primate brain with phylogenetic proximity to humans. Metabolism in the brain is critical for neurological function, as it provides cellular energy and critical biomolecules necessary for the complex cellular network characteristic of the brain [69-71, 157-160]. Cellular metabolism involves the breakdown of fuel molecules to produce energy or other molecules through multiple interconnected pathways, including glycolysis, oxidative phosphorylation, and the pentose phosphate pathway. Enrichments for metabolic processes in genes and gene regulatory regions undergoing positive selection is a common thread in gene expression analyses from whole primate brain tissue [39-42, 161, 162]. However, there are lineage-dependent differences in the specific pathways enriched between species (e.g. glucose and carbohydrate metabolism in humans and glycogen and acyl-CoA metabolism in chimpanzees) [42, 161, 162], as well regions of accelerated genomic change in a lineage-dependent manner [163, 164]. Within anthropoids, genes encoding the subunits of cytochrome c oxidase, the final component of the electron transport chain, show an accelerated rate of evolution in their sequences compared with any other placental mammals [165]. These changes at the gene level are indicative of increased control over the mechanisms that process glucose [162, 166-169]. Further understanding gene-phenotype relationships between genetic changes (both in coding and non-coding regulatory portions of the genome) and observed metabolic differences in primates will contribute to a greater understanding of proximate influences on larger evolutionary trends in primates. These findings also highlight a need to investigate not only glucose metabolism and energy production but also that of other macromolecules (e.g. lipids, amino acids, nucleic acids) for a more comprehensive understanding of differences in cellular metabolism in neural cells of primates.

The findings of interspecies divergence in brain metabolism are intriguing, however, a cell-type specific comparison would more fully inform our understanding of distinct cellular contributions to interspecific differences in neurological function [170]. Two of the major cell

types in the brain are neurons and astrocytes. Neurons are responsible for neurological functions like cognition and perception largely by transmitting chemical and electrical signals throughout complex cellular networks. However, non-dividing, mature neurons are known to have very little capacity for specific metabolic processes (e.g. glycolysis) and rely on metabolite shuttling from another cell type, astrocytes [171-173]. Astrocytes, despite being the most abundant cell-type in the central nervous system [174], have traditionally been considered support cells for neurons, without significant relevance to neural function. However, recent work has determined critical roles of astrocytes in neural function including provisioning of metabolites to neurons for energy [73, 175, 176] and enhancing synaptic processes [177, 178]. These findings point to a need to characterize the important differences among a variety of cell types, not only in neurons, but in other metabolically-relevant brain cell types such as astrocytes between species to understand how the primate brain has evolved metabolically.

We hypothesize that there are important cell-type specific metabolic changes between human and chimpanzee brains and that astrocytes contribute, at least in part, to these differences. To investigate these changes, we used established protocols for the differentiation of induced pluripotent stem cells (iPSCs) into mature, functional neurons and astrocytes from humans and chimpanzees from multipotent neural progenitor cells (NPCs). This novel comparative approach allowed us to functionally test each cell type in the absence of other cell types in a defined, controlled environment. In order to determine adaptive interspecies differences in gene expression and metabolism in cell-type specific manner, we conducted RNA-Seq on human and chimpanzee NPCs, neurons, and astrocytes.

3.3 Materials and Methods

3.3.1 Samples and Cell Culture

Induced pluripotent stem cells (iPSCs) from three individuals (cell lines) per species (human and chimpanzee) were cultured in defined, iPSC-specific media mTeSR1 (STEMCELL, Vancouver, Canada). These cell lines were originally obtained as fibroblasts from minimally invasive skin biopsies, reprogrammed into iPSCs, and have been extensively validated for their pluripotency and differentiation abilities [170, 179-184]. Three cell lines per species, representing three male individuals, were used (SI Table 3.1). To investigate differences between human and chimpanzee neural cell types, we induced iPSCs from each species first into multipotent, neural-lineage committed neural progenitor cells (NPCs) using STEMdiff Neural Induction Medium in monolayer for three passages (21-28 days), as per manufacturer's instructions (STEMCELL Technologies, Vancouver, Canada). Successful transition of iPSCs into NPCs was determined using immunofluorescence for the absence of the stem-cell marker OCT4 and presence of the NPC-marker PAX6 (Figure 7A). NPCs were then expanded into three subsets: one for RNA collection, and two for further differentiation and maturation into neurons and astrocytes. We then differentiated NPCs into mature neurons and astrocytes using the neuron and astrocyte specific STEMdiff differentiation and maturation kits as recommend by the manufacturer. Briefly, we differentiated NPCs using the STEMdiff Neuron Differentiation Medium for one week and then matured them using the STEMdiff Neuron Maturation Medium for two weeks. Similarly, we differentiated NPCs using the STEMdiff Astrocyte Differentiation Medium for three weeks and then matured them using the STEMdiff Astrocyte Maturation Medium for two weeks. All cells were validated for cell type via immunofluorescence prior to harvesting as follows: NPCs for PAX6+/OCT4- (Developmental Studies Hybridoma Bank, University of Iowa), neurons for neuron-specific class III β-tubulin (TUJ1; Neuromics), and astrocytes for GFAP (Sigma Aldrich), according to manufacturer's suggestions (SI Figure 13). All mature iPSC-derived cells for each cell type were harvested at similar timepoints: NPCs at passage 5-6 post-induction from iPSCs, mature neurons at passage 3-4 and mature astrocytes at passage 5-6 post-differentiation from NPCs and subsequent

maturation (SI Table 3.1). We used edgeR [58] to normalize our raw counts across all samples and visualized these data using a multidimensional scaling (MDS) plot of all of the expressed genes (Figure 7B).

3.3.2 Library Preparation and Sequencing

Total RNA was extracted from cells (1-2 wells, 6 well plate) using an RNeasy Plus Mini Kit (Qiagen), including a DNase step to remove residual DNA. Total RNA was analyzed for quality using the Agilent Bioanalyzer system (Agilent RNA 6000 Nano kit) with RNA Integrity Numbers (RINs) for all samples between 8.3-10 (SI Table 3.1). Using the NEBNext Poly(A) Magnetic mRNA Isolation Kit (NEB), mRNA was isolated from intact total RNA, and cDNA libraries were made from each sample using the NEBNext RNA Ultra II Library Prep Kit for Illumina (New England Biolabs). Barcoded samples were sequenced using the Illumina NextSeq 500 platform at the Genomics Resource Core Facility (Institute for Applied Life Sciences, UMass Amherst) to produce 75 base pair single-end reads, yielding a minimum of 32 million reads per sample.

3.3.3 Read Mapping and Quantification

Quality-filtered reads were aligned to respective species' most recent ENSEMBL genome (*Homo sapiens* GRCh38 and *Pan troglodytes* PanTro3.0 [53, 54]) with Bowtie2 [52] using default '--local' parameters for gapped alignments, with a minimum alignment percentage of $\geq$ 98.84% (SI Table 1). HT-Seq [55] was used to quantify counts per gene for each sample, using ENSEMBL gene transfer files (GTFs) corresponding to the same genome build used for alignment [56]. High quality, one-to-one orthologs from *P. troglodytes* were matched to the ENSEMBL human reference set of genes using biomaRt [57], yielding 15,284 genes identified as expressed in at least one sample.

3.3.4 Clustering Analyses

We used clustering analyses to determine the variation among our iPSC-derived samples as well as in comparison to previously published, publicly available data from other

tissues and cell types. For our iPSC-derived samples, we used the R package edgeR [58] to filter out lowly-expressed genes (counts per million (CPM) > 1 in 12/17 samples), resulting in 10,715 orthologous genes, and produced an MDS plot of our samples (Figure 7B). The greatest influence on our samples is species along PC1 and PC2, followed by separation of immature NPCs cells from mature cell types (neurons and astrocytes) along PC2 (Figure 7B). Notably, human samples were more variable than chimpanzee samples. One human cell line (H20961) showed significant variation across all cell types (SI Figures 14-16), however, the H20961 NPC sample was consistently an outlier, grouping outside of NPCs of either species, and was removed from subsequent analyses. There are no overt technical differences influencing this out-grouping (e.g. individual sex or age, RNA or cDNA library quality, read number, alignment percentages, SI Table 3.1). This cell line has successfully been used in before in other differentiation studies with no overtly different characteristics [170, 179-184].

To compare our samples to previously published data from cells and tissues, we downloaded raw RNA-Seq reads from the NCBI's Gene Expression Omnibus (GEO) [185] and processed them from raw read counts through HT-Seq and orthologous gene matching in the same manner as our iPSC-derived samples. To compare our samples to those from primary neural cell types, we used RNA-Seq data from primary neurons and astrocytes obtained from four hippocampal astrocytes, four cortex astrocytes, and one cortical neuron from [186] (GEO accession number GSE73721) and three pyramidal neuron samples (GEO accession numbers GSM2071331, GSM2071332, and GSM2071418) isolated from an unspecified brain region by the ENCODE project [187, 188]. We also downloaded the tissue-level data from Brawand et al., (2011) [45] from human and chimpanzee brain regions and non-neuronal tissue (heart, kidney, liver) (GEO accession number GSE30352) (SI Table 3.2 for details). Only genes with counts greater than zero in all samples were included and were further filtered to include only those with CPM > 1 in all 23 samples. An MDS plot of normalized counts was generated using edgeR of the top 500 most differentially expressed genes in all samples (SI Figure 17).

3.3.5 Differential Gene Expression Analyses

In order to determine what genes were significantly differentially expressed in a species by cell type manner using, we used the R package edgeR's [58] generalize linear model (GLM) functionality with a design matrix accounting for an interaction between species (SP) and cell type (CT) (referred to as SPxCT DE analysis). We performed an analysis of variance (ANOVA)-like test for differences across all samples. Furthermore, in order to determine what differences existed between species for each cell type, we performed interspecies pairwise DE comparisons in a similar manner between NPCs, neurons, and astrocytes (referred to as CT-DE analyses). We also used the GLM for these analyses, but did not include more than one cell type in these analyses in order to include genes that may be cell-type specific. For all analyses, we used edgeR's quasi-likelihood F-test and considered gene expression significantly different at a false discovery rate (FDR) of less than 5%. Normalization of data in edgeR for DE analyses ensured that DE is not dependent on original number of cells. All Venn diagrams were created using the R package Vennerable.

3.3.6 Categorical Enrichment Analyses

Uninformed pathway enrichment analyses were conducted using genes identified as differentially expressed from each DE comparison using gProfiler [189] with their functional enrichment tool (g:GOSt). Categorical enrichment analyses for overrepresented (enriched) and underrepresented (conserved) processes were conducted on all genes identified as differentially expressed (FDR < .05%) between species for individual cell types. Enrichments with a q-value of < .05 were considered significant.

3.3.7 Gene Set Enrichment Analyses

In order to investigate which metabolic pathways were enriched in a species' CT, we used Gene Set Enrichment Analyses (GSEA) [190]. We tested for enrichment of 23 *a priori* gene sets from the Molecular Signatures Database (MSigDB) [191] using the raw counts of the same set of genes used for the CT-DE pairwise comparisons. Gene sets were considered

significantly enriched according to suggested thresholds (FDR < 25% and nominal p-value < .05) [190]. Leading edge analyses determined a set of core enriched genes that most significantly influenced the enrichment of the gene set per phenotype.

3.3.8 Selection Analyses

In order to determine if genes exhibiting significant interspecies differential expression also had evidence of positive selection in their coding sequences, we used nonsynonymous (dN) and synonymous (dS) nucleotide changes per gene for all genes expressed in iPSC-derived neural cells. These were obtained from Ensembl using biomaRt (Kinsella *et al.*, 2011). These pre-calculated dN and dS values were originally computed by Ensembl using codeml and yn00 of the PAML package to compute dN and dS scores for each species in comparison to human (Herrero *et al.*, 2016). A rate of change was calculated for each gene (dN/dS), where a dN/dS > 1 is indicative of positive selection [192].

3.3.9 Network Schematic

We constructed a focal set of signaling pathways based upon HumanCyc [193] in order to contextualize our DE results in the framework of a network signaling, and this is the diagram of the major pathways involved in aerobic glycolysis (glycolysis, pentose phosphate pathway (PPP), lactate conversion from pyruvate, and TCA cycle) shown in Figure 12. For each enzyme in the pathway, three blocks indicate expression of this enzyme in each cell type – left to right: NPCs, neurons, astrocytes. Color indicates level of expression (higher in human (red), higher in chimpanzee (blue), not expressed in this cell type (grey)).

3.4 Results

3.4.1 RNA-Seq of human and chimpanzee iPSC-derived neural cells

We took a comparative genomics approach to investigating interspecies differences in neural cell-type specific gene expression between humans and chimpanzees. Three cell lines per species, representing three individuals, were used. These cell lines were originally obtained as

fibroblasts from minimally invasive skin biopsies, reprogrammed into iPSCs, and have been validated for their pluripotency and differentiation abilities [170, 179-184]. iPSCs from both species were initially cultured in the defined, iPSC-specific media mTeSR1 (STEMCELL, Vancouver, Canada). In order to investigate interspecies differences in cell-type specific gene expression between humans and chimpanzees, we generated RNA-Seq data from human and chimpanzee neural progenitor cells (NPCs), neurons, and astrocytes from induced pluripotent stem cells (iPSCs).

To confirm that expression profiles of our iPSC-derived neural samples resembled that of neural tissue and primary neural cells, we used previously published data from human and chimpanzee tissues, including brain [45], as well as human primary neurons and astrocytes (Methods) [186]. We visualized these data in an MDS plot and observed that our iPSC-derived neural samples clustered together within the same dimensional space as the other neural tissue and cell samples and not nearer to the non-neuronal tissue samples (SI Figure 17). These clustering analyses demonstrate that we successfully created and obtained total transcriptome data of iPSC-derived NPCs, neurons, and astrocytes from humans and chimpanzees relevant for comparative assessments of cell-type specific interspecies gene expression differences.

3.4.2 Differential expression between human and chimpanzee neural cell types

We next performed differential expression (DE) analyses in order to determine significantly differentially expressed genes between species. However, given the lack of clear distinction among our different cell types (Figure 7B), we first wanted to determine the degree of shared expression across all cell types to inform our investigation of cell-type specific differential expression. To do so, we determined overlap among cell types for genes with at least one count in one or more cell lines per cell type (Figure 7C). Of the total genes expressed in NPC (n=14,877), neuron (n=14,961), and astrocyte (n=14,931) samples, 95.13% (n=14,536) were shared among all three cell types (Figure 7C). This demonstrates that very few genes are

cell-type specific in the brain, similarly to other studies of gene expression across brain regions and cell types [72, 194].

For DE analyses, we first conducted an analysis of variance (ANOVA)-like test for differentially expressed genes in a species (SP) by cell-type (CT) manner using edgeR [58]. We reasoned that this would be the most evolutionarily relevant set of genes for investigating neural cell-type specific "trade-offs" in expression between species. Using edgeR's generalized linear model (GLM) functionality and a quasi-likelihood F-test for significant differential expression, we found 4,007 significantly differentially expressed genes in a species by cell type manner (26.22% of all expressed genes). However, at present, there are no post-hoc tests for an ANOVA-like test for differential expression, and so this analysis is limited in that it cannot delineate which samples (cell types) these genes are significantly DE in [58]. The ANOVA-like test for differences also requires an initial filtering of lowly expressed genes across all samples, which eliminates the 104-120 genes (0.68-0.79%, Figure 7C) expressed only in one cell type (CT-specific genes). For these reasons we also conducted interspecies pairwise DE comparisons for each CT (hereafter referred to as CT-DE analyses). While these CT-specific genes are relatively few in number, they likely have an important role in cellular function, and thus we did not want to exclude them from our interspecies CT-DE comparisons.

For CT-DE comparisons, the only genes included were those counts above zero in all samples per CT and were further filtered to those with counts per million (CPM) > 1 in at least 1 sample, resulting in 11,772 genes in NPCs, 12,451 genes in neurons, and 12,302 genes expressed in astrocytes. We used the same GLM quasi-likelihood F-test to determine that 8.57% (n=1,294) of genes are differentially expressed between species' NPCs, 5.8% (n=886) between neurons, and 12.2% (n=1,865) between astrocytes (Figure 8A, SI Figure 18). Many of these significantly differentially expressed genes in CT-DE comparisons overlapped with the SPxCT ANOVA-like differentially expressed genes (SI Table 3.3). When we determined overlap in differentially expressed genes between species across all three cell types, we found that, similar

to global expression, a large number of genes were determined as differentially expressed between species in all three cell types (n=594, Figure 8B). However, there are far more genes that uniquely differentiate astrocyte gene expression between species (n=924) than NPCs (n=395) and neurons (n=100) (Figure 8B). This suggests that neuronal gene expression is more conserved across species in NPCs and neurons, and that astrocytes do indeed contribute to important interspecies differences in neural gene expression.

3.4.3 Interspecies differences in gene expression are largely due to differential metabolic signaling skewed toward higher expression in humans, regardless of cell type

We then used categorical enrichment analyses to determine what biological processes are over-represented (enriched) or under-represented ("conserved") in interspecies differentially expressed genes by cell type (CT) (SI Table 3.4). There were consistently a larger number of interspecies differentially expressed genes per CT-DE comparison with higher expression in human cells (765 in NPCs, 496 in neurons, and 1,078 in astrocytes) than chimpanzee cells (529 in NPCs, 390 in neurons, and 787 in astrocytes) (Figure 8A). We used these six higher-in-one-species split DE gene lists in a mutliquery categorical enrichment analyses for under- and over-represented processes using gProfiler's categorical enrichment tool (gOST) [189]. Likely in part due to the larger number of genes with higher expression in human for all CT's, there was consistently far more processes enriched in human CTs than chimpanzee CTs (SI Table 3.4).

Human and chimpanzee neural cells exhibited significant under-representation of the KEGG pathways 'olfactory transduction' and 'neuroactive ligand-receptor interaction' (Figure 8C). Consistently, both species' also showed significant under-representation of GO BP related to development, immune function, and intracellular signaling (SI Figure 20). Human cells exhibited under-represented for some extracellular and membrane associated cellular components (CC) (SI Figure 21) as well as molecular function related to cytokine and receptor activity (primarily in astrocytes; SI Figure 22). Both species cells were underrepresented for nucleic acid binding and G-protein coupled receptor and transducer activity (SI Figure 22). This

demonstrates that signaling, including some neuronal-specific signaling such as neuroactive ligand receptor interaction, and downstream perception processes (e.g. olfaction) are conserved across species for all cell-types.

As for significantly over-represented processes in differentially expressed genes between species, cell division, cytoskeletal, developmental, signaling and response to external stimuli terms were significantly over-represented in human astrocytes, transcription was enriched in human NPCs, and protein modification was enriched most significantly in human neurons (SI Figure 20). Metabolic processes targeting a variety of substrates or macromolecules were enriched in human cells (summarized in Figure 8D, full results in SI Table 3.4). Human neurons and astrocytes were enriched for protein metabolic processes, while human NPCs were not. Conversely, human NPCs and astrocytes were enriched for metabolism of nitrogen containing and aromatic compounds while human neurons were not. Human astrocytes were significantly enriched for several more metabolic biological processes than human NPCs and neurons that included metabolism of phosphate-containing compounds as well as more generally for metabolism of macromolecules (Figure 8D). These results indicate that though interspecies divergence in metabolic gene expression does occur in all three cell types, it does not always do so for the same metabolic pathways in all cell types.

We also investigated enrichment of GO cellular component (CC) terms to determine if there were differences in expression of specific neuronal parts. Over-represented CC terms in human astrocytes were similar to the GO BP over-represented processes (cytoplasm, cytoskeleton; (SI Figure 21A) as well as cell-parts involved in cellular division and growth (SI Figure 21B). Human neurons were enriched for terms related to the Golgi apparatus and macromolecule complexes (SI Figure 21A) and those specific to protein modification and signaling (SI Figure 21B). Human astrocytes were enriched for molecular functions (MFs) related generally to substrate binding, and specifically, to ATP, carbohydrates & their derivatives, enzymes, and nucleic acids (SI Figure 22A). The only over-represented GO MF

binding activity for chimpanzee cells was for that of oxidoreductase activity in chimpanzee NPCs (SI Figure 22B). Human NPCs were enriched for TF activity, human neurons for ubiquitin-related molecular activity, and human astrocytes for molecular activity related generally to ATPases, catalysis, exonucleases, helicases, kinases, and phosphotransferases (SI Figure 22B). These results indicate that metabolic processes differ between species in a CT-specific manner, and that all human neural cell types exhibit increased expression for a variety of macromolecular metabolic processes more than chimpanzee neural cells. Further, we see that there are significant differences in molecular functions important in cellular metabolic signaling.

3.4.4 Human and chimpanzee neural cells differ in glucose and lactate transport as well as oxidative phosphorylation

Metabolic processes targeting a variety of substrates or macromolecules were enriched in human neural cell types using uninformed categorical enrichment analyses. However, very few of these processes were for pathways involved in cellular respiration resulting in production of energy (in the form of ATP). We were interested if there were significant interspecies differences in gene expression for pathways involved in cellular respiration, specifically, those involved in aerobic glycolysis. Because there are known differences in metabolic capacity between neurons and astrocytes, including that astrocytes are characterized metabolically by high aerobic glycolytic activity (increased glycolysis and limited potential for oxidative ATP production) while neurons typically favor energy production and oxidative phosphorylation [reviewed in 72], we were interested in determining any interspecies, cell-type specific differences in these brain metabolic processes. To further investigate this, we used a Gene Set Enrichment Analysis (GSEA) [190] with 23 *a priori* gene sets on the raw counts of the 12,407 genes used for interspecies pairwise CT-DE analyses. Gene sets were obtained from the Molecular Signatures Database (MSigDB) [191] and chosen in order to probe a variety of energetic metabolic pathways and substrate transporters of varying gene number size from

multiple ontology categories (GO, KEGG, and REACTOME) (all probed gene sets listed in Figure 9A) [64-66, 195]. The goal was to determine if pathways involved in aerobic glycolysis (e.g. oxidative phosphorylation, glucose transport, TCA cycle) differ in a species by cell type manner, and so pathways not directly involved in aerobic glycolysis (e.g. fatty acid metabolism) are included as a comparison. We also included "control" pathways not directly related to metabolism (regulation of growth, and neurotrophin signaling).

Our GSEA results indicate that the gene sets for lactate transmembrane transport, glucose transport, oxidative phosphorylation, and metabolism of RNA are significantly different between human and chimpanzee neural cells (FDR < 25% and nominal p-value < .05; Figure 9A, B, SI Table 3.5). Glucose transport was enriched in human neurons and astrocytes while lactate transmembrane transport was enriched in chimpanzee neurons and astrocytes (Figure 9A, B). The GO gene set for oxidative phosphorylation is significantly enriched in all human cell types while the KEGG oxidative phosphorylation is upregulated in chimpanzee astrocytes (Figure 9A, B). These results indicate cell-type by species differences in glucose uptake by cells, lactate shuttling, and diverging energetic cellular respiration.

3.4.5 There are species by cell-type differences in expression of oxidative phosphorylation protein complexes

Leading edge analyses of significant GSEA gene sets are used to determine which genes of the gene set contribute most strongly to the enrichment of that pathway in the phenotype [190]. We examined the results from the leading edge GSEA analysis with CT-DE expression analyses to get a better idea of how these three pathways diverge in a cell type by species manner (Figures 10 and 11; full results in SI Table 3.6). We calculated a rank for DE genes for each CT-DE comparison (NPC, neuron, and astrocyte): (sign of logFC) x log10(FDR Q-value) [196] and used that in addition to the GSEA leading edge analysis to determine significant differences. For the oxidative phosphorylation genes, we were interested in determining why there was a difference in species and cell type enrichment based on the source

of the gene set (KEGG vs GO) and determining if there were potential functional differences in oxidative phosphorylation between species. We mapped the CT-DE rank of the core enriched genes for GO (Figure 10A) and KEGG (Figure 10B) oxidative phosphorylation genes.

The core set of genes in GO and KEGG oxidative phosphorylation (OXPHOS) gene sets included genes for subunits of cytochrome c oxidase (the nuclear-encoded *COX4I1*, *COX6B2*, *COX7B*, and *COX7C* and the mitochondrially-encoded *MT-CO1*, *MT-CO2*, and *MT-CO3*) as well as those that aid in cytochrome c oxidase assembly (*COX10* and *COX11*) (Figure 10A, B; SI Table 3.6). Cytochrome C oxidase is the terminal complex in the electron transport chain and is crucial to maintaining a proton gradient across the inner mitochondrial membrane for ATPase to synthesize ATP. These genes are of particular interest, because within anthropoids, genes encoding the subunits of cytochrome c oxidase show an accelerated rate of evolution in their sequences compared with any other placental mammals [165]. Here, we see a cell type by species divergence in cytochrome c oxidase gene expression, where most of these genes exhibit higher expression in human neurons (Figure 10A). There is also a clear trend of mitochondrially-encoded genes that function in OXPHOS having significantly higher expression in human cells, including those for cytochrome c oxidase subunits (*MT-CO1*, *MT-CO2*, *MT-CO3*), but also mitochondrially-encoded ATP synthase (*MT-ATP6*) and mitochondrially-encoded subunits of the NADH:ubiquinone oxidoreductase core of electron transport chain complex I (*MT-ND2*, *MT-ND3*, *MT-ND5*, and *MT-ND6*) (Figure 10A).

A major difference between the GO and KEGG OXPHOS gene sets is that the KEGG OXPHOS includes vacuolar-ATPase (V-ATPase) genes, whose major role is in acidification of intracellular organelles, and have an important function in synaptic vesicle proton gradient formation and maintenance [197, 198]. There is an intriguing pattern of enrichment for higher expression of subunits of V-ATPases in a cell type by species manner (Figure 10). Three genes for subunits of vacuolar ATPases (*ATP6V0C*, *ATP6V0D1*, and *ATP6V1C2*) are core enriched genes in the KEGG OXPHOS gene set and are significantly enriched in CT-DE with higher

expression in chimpanzee NPCs and astrocytes, but not neurons (Figure 10B). Furthermore, only one of these vacuolar-ATPase genes is DE in a SPxCT manner (*ATP6V1C2*) (Figure 10B). However, several other V-ATPase subunit genes are core enriched only in human neurons (Figure 10B), most notably *ATPV1E2* and *ATPV0E2*, both of which are core enriched and significantly differentially expressed in human neurons. This shows that V-ATPases exhibit significant DE between humans and chimpanzees, and that human neurons are distinct in V-ATPase gene expression from chimpanzee NPCs and astrocytes. Given the important function of V-ATPases in synaptic vesicle formation for neurotransmitter signaling, this may imply an important functional change in humans specifically in neurons.

3.4.6 Interspecies differential expression of important metabolite transporter genes in neurons and astrocytes

It is widely accepted that neurons exhibit limited glycolytic capacity and that astrocytes respond to signals associated with increased synaptic signaling by increasing glucose uptake and subsequent aerobic glycolysis of that glucose to lactate, to be used as energy source necessary by neurons [reviewed in 72, 175]. For this reason, we were interested in investigating if there were interspecies gene expression differences in lactate transport, particularly in neurons and astrocytes. The GO gene set 'lactate transmembrane transport' was enriched in chimpanzee neurons and astrocytes, showing that genes involved in lactate transport are more highly expressed in these mature cell types than NPCs (Figure 11A). Several genes in this gene set are for proton-linked monocarboxylate transporters that transport pyruvate and lactate (*SLC16A11*, *SLC16A12*, *SLC16A13*, and *SLC16A6*) that are all core enriched in neurons and astrocytes (Figure 11A). *SLC6A11* and *SLC16A13* are also differentially expressed in SPxCT ANOVA-like DE as well as an CT-DE manner, though *SLC16A13* is not in astrocytes (Figure 11A). The enrichment for the lactate transmembrane transport gene set in chimpanzee neurons and astrocytes and the corresponding DE of specific pyruvate and transporter genes between

species' suggest that chimpanzee neurons and astrocytes have the capacity to shuttle pyruvate and lactate at a higher rate than human neurons and astrocytes.

In addition to lactate transmembrane transport enrichments, the glucose transport gene set was significantly enriched in human neurons and astrocytes (Figure 9). There were two hexokinase genes (*HK1* and *HK2*) core enriched in this gene set that demonstrate lower expression in chimpanzee NPCs but higher expression in human astrocytes (Figure 11B), though only *HK2* is significantly upregulated in human astrocytes by CT-DE analysis (Figure 11B). *G6PC3* may not be significantly differentially expressed in any particular cell type, but it is in the SPxCT DE comparison and does show insignificant but consistently higher expression in all chimpanzee cell types (Figure 11B). *SLC2A3* is a facilitative glucose transporter across the cell membrane, and here, it exhibits core enrichment in all three human cell types by the GSEA leading edge analysis, as well as moderately (though non-significant) higher expression in human (Figure 11B). The enrichments of glucose transport in human neurons and astrocytes appears to be influences by increased expression of plasma membrane associated glucose transporters (e.g. *SLC2A3*) and enzymes that function in the earlier steps of glycolysis (*HK1, HK2, G6PC3*).

3.4.7 Few genes exhibiting interspecies DE in iPSC-derived neural cells have signs of positive selection in their coding regions

In order to begin to probe whether expression differences between species are influenced by selective pressures, we obtained synonymous (dS) and nonsynonymous (dN) nucleotide mutation rates from Ensembl [53, 54] and compared the rate of change (dN/dS) for different groups of iPSC-derived neural cell expressed genes. A dN/dS > 1 indicates putative evidence of positive selection in coding regions [192]. As predicted, the vast majority of all the genes identified as expressed in these cells did not exhibit a dN/dS > 1 (SI Figure 23). Only a few genes DE between species in iPSC-derived astrocytes (n=6), neurons (n=6), and NPCs (n=11) exhibit signs of coding selection (dN/dS > 1) (Table 5, SI Figure 23). The gene *HRC*

(histidine rich calcium binding protein) exhibits positive selection and is significantly DE between species in NPCs, but not astrocytes, and is not expressed at all in neurons (Table 5). Three genes (*DCTN6*, *HHLA3*, *DBNDD2*) have a dN/dS > 1and showed significantly DE in all three cell types (Table 5). However, there is no commonality in these genes to suggest any meaningful impact on gene expression differences or in specific cell types. This is likely due, in part, to the limitations of the methods used for probing coding sequences, rather than those in non-coding, *cis*-regulatory changes, which previous studies have demonstrated are critical for significant differences in expression in primate brains [40-42].

3.4.8 Significant interspecies differences in the expression of aerobic glycolysis enzymes are primarily in NPCs and astrocytes but not neurons

In order to obtain a pathway-level understanding of altered expression of aerobic glycolysis in iPSC-derived neural cells between humans and chimpanzees, we reconstructed a signaling network diagram of enzymes involved in four sub-pathways involved in aerobic glycolysis (glycolysis, pentose phosphate pathway, pyruvate conversion to lactate, and the citric acid (TCA) cycle) from the HumanCyC database (Figure 12) [193]. We then mapped discrete expression values (higher in human, higher in chimpanzee, not expressed) for each of these enzymes in all three cell types onto the pathway diagram to illustrate which species the enzymes were more highly expressed in and if they were significantly differentially expressed between species in which cell types. From this, we see dynamic changes in expression across these aerobic glycolysis sub-pathways, with no significant shift towards higher expression of enzymes in one species or cell type at any of these sub-pathways (Figure 12). Aerobic glycolysis enzymes exhibiting interspecies DE in NPCs were *PGD*, *LDHA*, *LDHB*, *LDHC*, and *ME1*, enzymes demonstrating DE in astrocytes were *PGD*, *HK2*, *PDHA1*, *FH*, *ACO2*, and *SDHD,* and only a single enzyme exhibited interspecies DE in neurons (*LDHB*) (Figure 12). This shows that NPCs and astrocytes, but not neurons, exhibit the vast majority of significant differences in expression of enzymes in these pathways (Figure 12). The vast majority of these genes were

expressed in all cell types, particularly those that exhibited significant DE in NPCs or astrocytes, so this lack of DE in neurons is not simply due to cell-type specific expression differences (SI Figure 24). Human and chimpanzee astrocytes appear to diverge at the stage of pyruvate utilization, where human astrocytes exhibit significantly higher expression of *PDHA1*, which converts pyruvate into acetyl-coA whereas chimpanzee astrocytes show significantly higher expression of *LDHB*, which converts pyruvate into lactate rather than acetyl-coA (Figure 12). Interestingly, *LDHB* is also the only enzyme in these pathways exhibiting differential expression between species in neurons. Other *LDH* isoforms (*LDHC* and *LDHA*) also exhibit significant interspecies DE, with higher expression in chimpanzee NPCs. All chimpanzee neural cell types differ from human chimpanzees for *LDH* expression, but NPCs differ from humans in expression levels of multiple *LDH* isoforms. This pathway level consideration expression differences between species suggests significant changes in aerobic glycolysis enzyme activity primarily in NPCs and astrocytes and an interspecies divergence in pyruvate utilization.

3.5 Discussion

This study represents, to our knowledge, the only investigation of neuronal cell-type specific differences in gene expression divergence across primate species. Our novel approach using iPSCs work allowed us to investigate rare neural cell types from primates to determine cell type specific metabolic changes necessary to support evolution of the human brain. Our results demonstrate interspecies divergence in gene expression is more conserved in neurons and significantly greater between species' astrocytes. Differential expression between species' cell types is enriched for metabolic processes – similarly to that of previous studies of differential expression between human and chimpanzee whole brain tissue – and is driven primarily by higher expression of metabolic genes in human cells [42, 161, 162]. We also determined that human neurons and astrocytes are enriched for higher expression of glucose transport proteins while chimpanzee neurons and astrocytes exhibit higher expression of lactate

transmembrane transport genes and that there are dynamic interspecies changes in expression of nuclear- and mitochondrially-encoded subunits of the protein complexes important for oxidative phosphorylation. Our study demonstrates the utility of iPSC-derived cells for better understanding evolution of gene expression in primates.

Though there were many important differences in metabolic expression between human and chimpanzee brains, the heterogenous nature of brain tissue complicated drawing specific conclusions about role of specific cell types on this trajectory of elevated metabolic expression in human brains. There have been long-standing questions about sole influence of greater neuron numbers in human brains [143] on the observed increase in glucose utilization [84, 85]. Our novel approach using iPSC-derived neural cell types and compairing gene expression differences in a cell type by species manner allowed us to test interspecies differences in the cell-type specific contribution of metabolism to long-understood differences in brain metabolic capacity. Because our methods of detrmining signficant differences in gene expression do not rely on number of cells or quantity of transcripts, we are able to conclude that there are cell-type specific contributions to altered metabolic gene expression between species' neural cell ttypes, and that sheer number of cells alone likely does not fully explain metabolic differences between human and chimpanzee brains.

Our results demonstrate that, in light of the relatively recent discovery of cell-type specific metabolic differences between neurons and astrocytes, investigation of differences in brain metabolism among primates and the evolutionary processes that shaped them would indeed be incomplete without the consideration of all metabolically-relevant neural cell types, not just neurons. We determined that astrocytes demonstrate the greatest proportion of interspecies difference in metabolic gene expression, and that, in fact, neuronal gene expression is more conserved across species. This suggests that astrocyte-mediated differences in metabolic brain function may be an important mechanism by which the ultimate evolutionary trajectory of human brain evolution has occurred.

Human cells show increased capacity for glucose transport, via greater expression of glucose transporters, into the mature neural cell types. This may suggest that the observed differences in glucose utilization by the human brain extend beyond development and may play an important role in more mature neurons and astrocytes for either energy or macromolecule production. Furthermore, lactose dehydrogenase (LDH) isoforms favor differential affinities for interconverting pyruvate and lactate. *LDHB* favors the production of lactate into pyruvate, and is significantly differentially expressed with higher expression in chimpanzee for all cell types [71, 199]. This, coupled with the increase in lactate transporters in chimpanzee astrocytes, suggests that chimpanzee cells may be favoring production and transport of lactate at a higher rate than human neural cell types. This opposing enrichment for elevated glucose transport in mature human neural cell types in comparison to elevated lactate transport and conversion to pyruvate raises some intriguing questions about metabolic trade-offs between human and chimpanzee brains. If we presume that the direction of change in metabolic gene expression is on the human lineage, and we do have some evidence from signs of positive selection on glucose and energetic metabolism coding and non-coding genes within primates with proximity to humans [162, 166-169], then has an increase in glucose uptake in human brains allowed for a coinciding decrease in gene expression focusing on converting and shuttling lactate (LDH and lactate transport) to produce pyruvate? Our current analysis of evidence of selection in genes exhibiting DE between species CTs did not reveal much insight. However, expanding this investigating into a more nuanced investigation of positive selection in non-coding regions of known function in altering expression levels of these interspecies DE genes would provide a more definitive assessment of the correlation between the functional changes we see here in aerobic glycolytic expression and the selective pressures that shaped it. Similarly, previous studies have found lineage-dependent differences in enrichments for metabolic pathways in genes DE between primate brain regions, where glucose and carbohydrate metabolism in elevated in humans and glycogen and acyl-CoA metabolism in chimpanzees [42, 161, 162]. The

increase in LDH and lactate transport in chimpanzee neurons and astrocytes may be a mechanism by which these previous findings of altered metabolic signaling at the whole tissue level occurred.

Human neural cells were enriched at the pathway level for oxidative phosphorylation genes, and within that pathway, there were some interesting examples of opposing enrichment for subunits of oxidative phosphorylation protein complexes. We not only saw increased expression and enrichment for components of cytochrome c oxidase, which previous studies have already determined genes involved in this complex to be under positive selection [166]. However, we expand on the knowledge of interspecies differences in cellular respiration complex expression by determining that these components are higher in human in all cell types investigated though there is a more consistent enrichment for higher expression of all of these genes in human neurons (Figure 10A). We also observed higher expression for subunits of other electron transport chain complexes, including ATP-synthase and the NADH:ubiquinone oxidoreductase components of complex I. This cell-type by species approach also allowed for us to determine that there are oxidative phosphorylation genes where human neurons and chimpanzee NPCs and astrocytes have higher expression for genes involved in vacuolar-ATPase function. Astrocytes respond to signals associated with increased synaptic signaling [reviewed in 72, 175] by increasing glucose uptake and subsequent aerobic glycolysis of that glucose to lactate, which is the energy source necessary for neurons. This suggests that V-ATPases are significantly differentially expressed between humans and chimpanzees, and that human neurons are distinct in V-ATPase gene expression from chimpanzee NPCs and astrocytes. Given the important function of V-ATPases in synaptic vesicle formation for neurotransmitter signaling, this may imply an important functional change in humans specifically in neurons.

Our investigation into the overlap of signatures of positive selection with genes exhibiting interspecies DE revealed very little new or intriguing information. This is likely due

at least in part to the limited scope of the current methods for searching for positive selection. The dN and dS scores obtained from Ensembl for use in this analysis were averages across all sites in a given gene, thus minimizing significant changes at specific sites [200]. More importantly, it is necessary to investigate positive selection in both coding and non-coding regulatory regions (i.e. promoters, *cis*-regulatory regions) associated with metabolic genes, as changes in both gene sequence and gene expression can confer adaptive phenotypic differences [39, 41, 42, 161, 162]. Non-coding regulatory regions of genes in oxidative phosphorylation and glycolysis pathways are found to be under positive selection to varying degrees in humans [39]. A specific example is glucose-6-phosphate isomerase (*GPI*), which has been determined to be under positive selection in its non-coding promoter region in previous studies [42], however, it does not exhibit interspecies DE in any of these cell types. Future investigations of the overlap between genes exhibiting DE in a cell-type specific manner and signatures of positive selection should utilize methods that allow for branch and site models that are more effective at determining positive selection in a lineage-specific manner (e.g. HyPhy for non-coding and coding regions) [42, 201-203].

Our focal analysis of aerobic glycolysis enzyme expression had a few important findings. First, we show that there is consistent a single species-skew in expression levels for any of the four sub-pathways in aerobic glycolysis. The lack of significant DE between species in neurons for aerobic glycolysis enzymes demonstrates the importance of studying cell types other than neurons for better understanding human brain evolution and suggests that astrocytes may indeed be critical for the evolution of human-specific brain gene expression. Secondly, we found that it is not simply the lack of glycolytic capacity of neurons [171-173] that contribute to cell-type specific signaling disparities, at least in a comparative manner. Lastly, perhaps the most intriguing finding is the interspecies divergence in processing pyruvate – humans exhibit significantly higher expression of *PDHA1* than chimpanzees do, indicative of a functionally relevant increase in conversion of pyruvate into acetyl-CoA and further utilization of the

products of glycolysis for energy production, while chimpanzee astrocytes exhibit expression phenotypes suggestive of greater lactate production (higher expression of *LDH*, which converts pyruvate to lactate) as well as enrichment for greater lactate transmembrane transport in chimpanzee neural cell types (higher expression of lactate transmembrane transporters). This suggests that chimpanzee neural cells, and most prominently astrocytes, have a significantly greater capacity to convert pyruvate into lactate and then shuttle it across membranes than human astrocytes do. These analyses suggest significant interspecies changes in aerobic glycolysis enzyme activity primarily in NPCs and astrocytes and an interspecies divergence in pyruvate utilization.

A potential drawback to a cell-type specific study of interspecies differences in neuronal gene expression is the loss of intercellular signaling between different cell types, a hallmark of synaptic signaling in whole tissue. Furthermore, the astrocyte-neuron lactate shuttle links the complementary metabolic needs of astrocytes and neurons [reviewed in 72, 175]. Neurons are considered largely to lack the ability to increase glycolytic activity. Future studies of gene expression differences with controlled levels of intercellular signaling by building in complexity (e.g. interspecies differences in expression of single cell types compared to that of co-cultured iPSC-derived neurons and astrocytes) could further inform interspecies differences in neuronal gene expression. We also determined cell-type by species differences in gene expression for nuclear- and mitochondrially-encoded subunits of the protein complexes important for oxidative phosphorylation. Finally, we show a significant interspecies divergence in aerobic glycolytic gene expression in astrocytes, suggesting that this traditionally understudied glial cell type likely contributes to the tissue-level shifts in gene expression and suggests that astrocytes play an important role in the evolution of the metabolically expensive human brain.

3.6 Conclusion

Evolved differences in metabolic investment may be the basis for a number of primate-specific phenotypes, including those that are uniquely-human, such as slow reproduction and growth and correspondingly longer lifespan than other placental mammals [151, 204, 205]. Adaptation can act on metabolic phenotypes through alterations to energy budget including reductions of or increases in total energy budget or differential allocation of energy within energy budget [151, 206, 207]. Our results demonstrate that altered gene expression between species' astrocytes - an under-studied, but critical brain cell type with known metabolic relevance, providing insight into the metabolic changes that were necessary to support evolution of the human brain.

3.7 Supplemental Information Table List

SI Table 3.1 | Sample, cell line, and sequencing information.

SI Table 3.2 | External RNA-Seq Metadata from Gene Expression Omnibus.

SI Table 3.3 | Lists of DE genes from edgeR.

SI Table 3.4 | Categorical enrichment analyses results from gProfiler.

SI Table 3.5 | GSEA Pathway enrichment table.

SI Table 3.6 | Table of leading edge analyses of genes in three significant GSEA pathways.

3.8 Main Figures and Tables

Figure 7

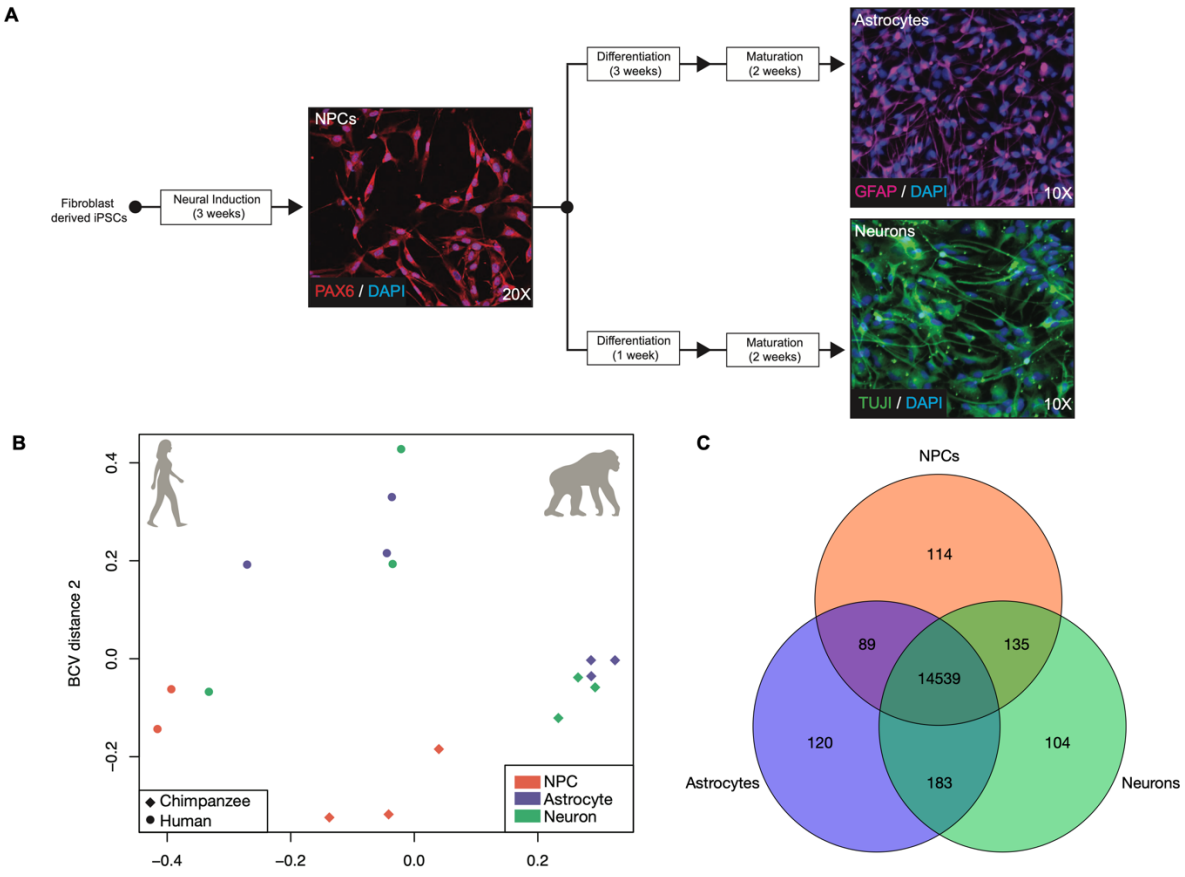


Figure 7 | Patterns of gene expression variation of iPSC-derived neural cells from humans and chimpanzees. A) The differentiation schematic and representative immunofluorescent photos of iPSC-derived neural progenitor cells (NPCs; line H28815) stained with PAX6, astrocytes (line C3649K) stained with GFAP, and neurons (line C4955) stained with TUJ1. B) A PCoA of the iPSC-derived NPCs, neurons, and astrocytes transcriptomes. C) A Venn diagram of the overlap in expression across cell types. Further details for samples are included in SI Table 1.

Figure 8

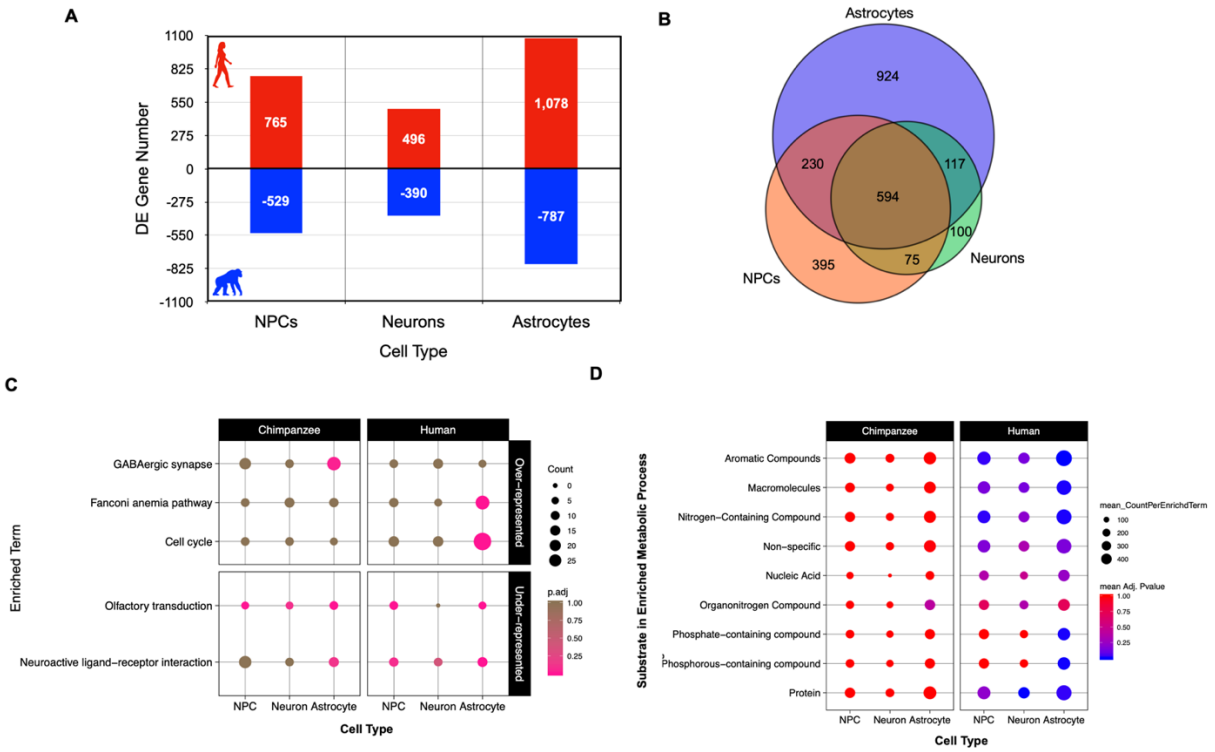

A
1100
825
550
275
0
-275
-550
-825
-1100
DE Gene Number
765
-529
496
-390
1,078
-787
NPCs
Neurons
Astrocytes
Cell Type
B
Astrocytes
924
230
117
594
100
395
75
NPCs
Neurons
C
Chimpanzee
Human
GABAergic synapse
Fanconi anemia pathway
Cell cycle
Olfactory transduction
Neuroactive ligand−receptor interaction
Over−represented
Under−represented
Enriched Term
Count
p.adj
NPC
Neuron
Astrocyte
Cell Type
D
Chimpanzee
Human
Aromatic Compounds
Macromolecules
Nitrogen−Containing Compound
Non−specific
Nucleic Acid
Organonitrogen Compound
Phosphate−containing compound
Phosphorous−containing compound
Protein
Substrate in Enriched Metabolic Process
mean_CountPerEnrichdTerm
mean Adj. Pvalue
NPC
Neuron
Astrocyte
Cell Type

Figure 8 | Astrocytes demonstrate the most significant differences in gene expression between human and chimpanzee neural cell types for metabolic but not neuron-specific pathways. A) Counts of genes exhibiting differential expression at an FDR < 5% between species for each cell type and the direction of higher expression for each CT-DE comparison (red/positive – higher expression in human; blue/negative – higher expression in chimpanzee). B) A Venn diagram of overlap in genes per cell type exhibiting differential expression between species. C) Plot of significantly (EASE < .05) over-represented (top panel) and under-represented (bottom panel) KEGG pathways determined by categorical enrichment analyses (full results in SI Table 4). Size indicates the count of genes per pathway while color indicates the adjusted p-value (pink – lower/significant, brown – higher/non-significant). D) Plot of significantly over-represented categories of GO BP terms determined by categorical enrichment analyses. (full results in SI Table 4). The categories (y-axis) represent groupings of multiple GO BP terms related to the metabolism of the indicated substrates/macromolecules. Size indicates the mean count and color indicates the mean adjusted enrichment p-value for all terms in that category. Refer to SI Table 4 and Methods for individual GO BP terms included in each category.

Figure 9

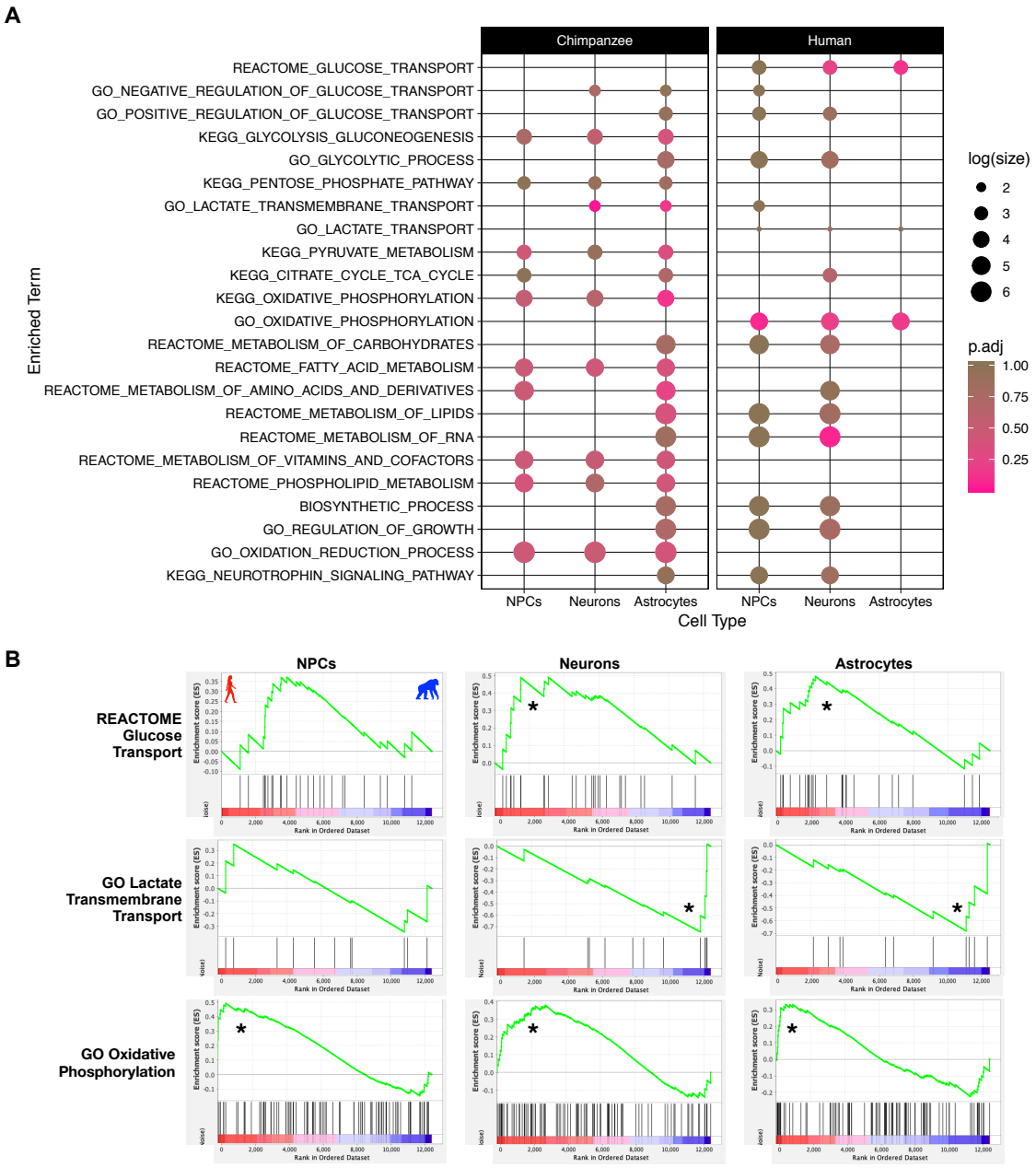

A
Chimpanzee
Human
REACTOME_GLUCOSE_TRANSPORT
GO_NEGATIVE_REGULATION_OF_GLUCOSE_TRANSPORT
GO_POSITIVE_REGULATION_OF_GLUCOSE_TRANSPORT
KEGG_GLYCOLYSIS_GLUCONEOGENESIS
GO_GLYCOLYTIC_PROCESS
KEGG_PENTOSE_PHOSPHATE_PATHWAY
GO_LACTATE_TRANSMEMBRANE_TRANSPORT
GO_LACTATE_TRANSPORT
KEGG_PYRUVATE_METABOLISM
KEGG_CITRATE_CYCLE_TCA_CYCLE
KEGG_OXIDATIVE_PHOSPHORYLATION
GO_OXIDATIVE_PHOSPHORYLATION
REACTOME_METABOLISM_OF_CARBOHYDRATES
REACTOME_FATTY_ACID_METABOLISM
REACTOME_METABOLISM_OF_AMINO_ACIDS_AND_DERIVATIVES
REACTOME_METABOLISM_OF_LIPIDS
REACTOME_METABOLISM_OF_RNA
REACTOME_METABOLISM_OF_VITAMINS_AND_COFACTORS
REACTOME_PHOSPHOLIPID_METABOLISM
BIOSYNTHETIC_PROCESS
GO_REGULATION_OF_GROWTH
GO_OXIDATION_REDUCTION_PROCESS
KEGG_NEUROTROPHIN_SIGNALING_PATHWAY
Enriched Term
NPCs
Neurons
Astrocytes
Cell Type
log(size)
2
3
4
5
6
p.adj
1.00
0.75
0.50
0.25
B
NPCs
Neurons
Astrocytes
REACTOME Glucose Transport
GO Lactate Transmembrane Transport
GO Oxidative Phosphorylation
Enrichment score (ES)
Rank in Ordered Dataset

Figure 9 | Humans and chimpanzees differ in metabolite transport and oxidative phosphorylation in a neural cell-type manner. A) Plot of all tested gene sets in the Gene Set Enrichment Analysis (GSEA) (full results in SI Table 5). Separate panels indicate which species 'phenotype' the gene set was enriched in. Color indicates the FDR Q-value (FDR < 25% indicates significance in this analysis). Size indicates the log(count) of genes included in the enriched gene set. B) Enrichment plots of significant pathways from GSEA for a subset of panel A. The green line indicates the running enrichment score for each gene in the gene set as the analysis moves down the ranked list of genes. The enrichment score for the gene set is the peak of this curve and an (*) indicates significantly enriched. The bottom panel is the ranked order of the genes and shows their location within that ranked set of genes. Left side of plot (and red/left portion of ranked order plot below) indicates human enrichment, while the opposite (right/blue) indicates chimpanzee enrichment. Full results in SI Table 5.

Figure 10

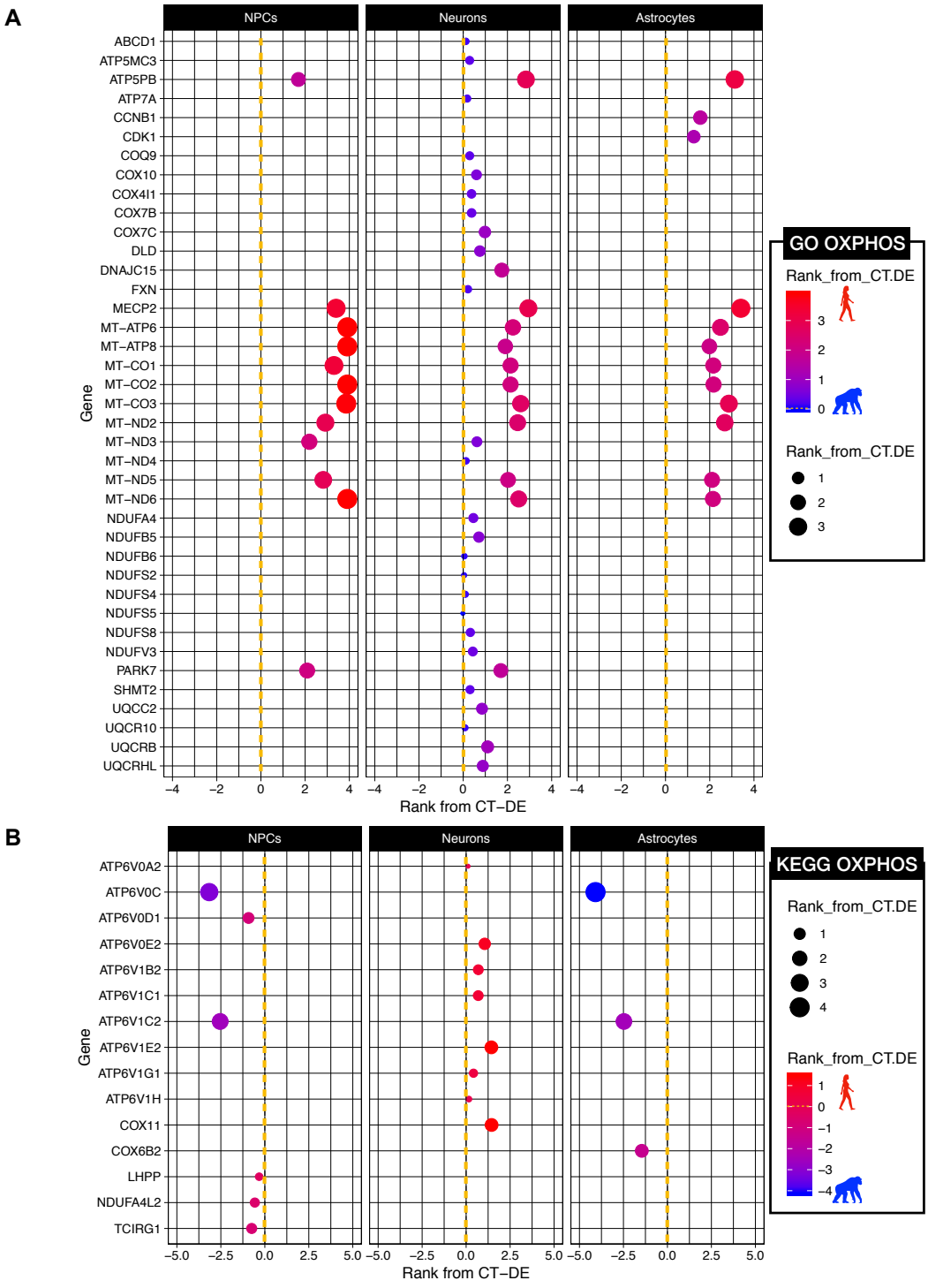

A
NPCs
Neurons
Astrocytes
ABCD1
ATP5MC3
ATP5PB
ATP7A
CCNB1
CDK1
COQ9
COX10
COX4I1
COX7B
COX7C
DLD
DNAJC15
FXN
MECP2
MT−ATP6
MT−ATP8
MT−CO1
MT−CO2
MT−CO3
MT−ND2
MT−ND3
MT−ND4
MT−ND5
MT−ND6
NDUFA4
NDUFB5
NDUFB6
NDUFS2
NDUFS4
NDUFS5
NDUFS8
NDUFV3
PARK7
SHMT2
UQCC2
UQCR10
UQCRB
UQCRHL
Gene
−4 −2 0 2 4
Rank from CT−DE
GO OXPHOS
Rank_from_CT.DE
3
2
1
0
Rank_from_CT.DE
1
2
3
B
NPCs
Neurons
Astrocytes
ATP6V0A2
ATP6V0C
ATP6V0D1
ATP6V0E2
ATP6V1B2
ATP6V1C1
ATP6V1C2
ATP6V1E2
ATP6V1G1
ATP6V1H
COX11
COX6B2
LHPP
NDUFA4L2
TCIRG1
Gene
−5.0 −2.5 0.0 2.5 5.0
Rank from CT−DE
KEGG OXPHOS
Rank_from_CT.DE
1
2
3
4
Rank_from_CT.DE
1
0
−1
−2
−3
−4

Figure 10 | Interspecies expression differences of oxidative phosphorylation genes is influenced by higher expression of mitochondrial genes in all neural human cell types. CT-DE results of genes per A) GO or B) KEGG oxidative phosphorylation gene sets determined as members of the core set genes influencing significant enrichment of these gene sets in the GSEA analysis. CT-DE rank per each gene (x-axis) - values greater than zero indicate higher expression in human and values less than zero indicate higher expression in chimpanzee. Color spectrum and size also indicate rank (red – higher in human, blue – higher in chimpanzee, larger = higher rank). See SI Table 3 for DE results per gene and SI Table 6 for GSEA results per gene.

Figure 11

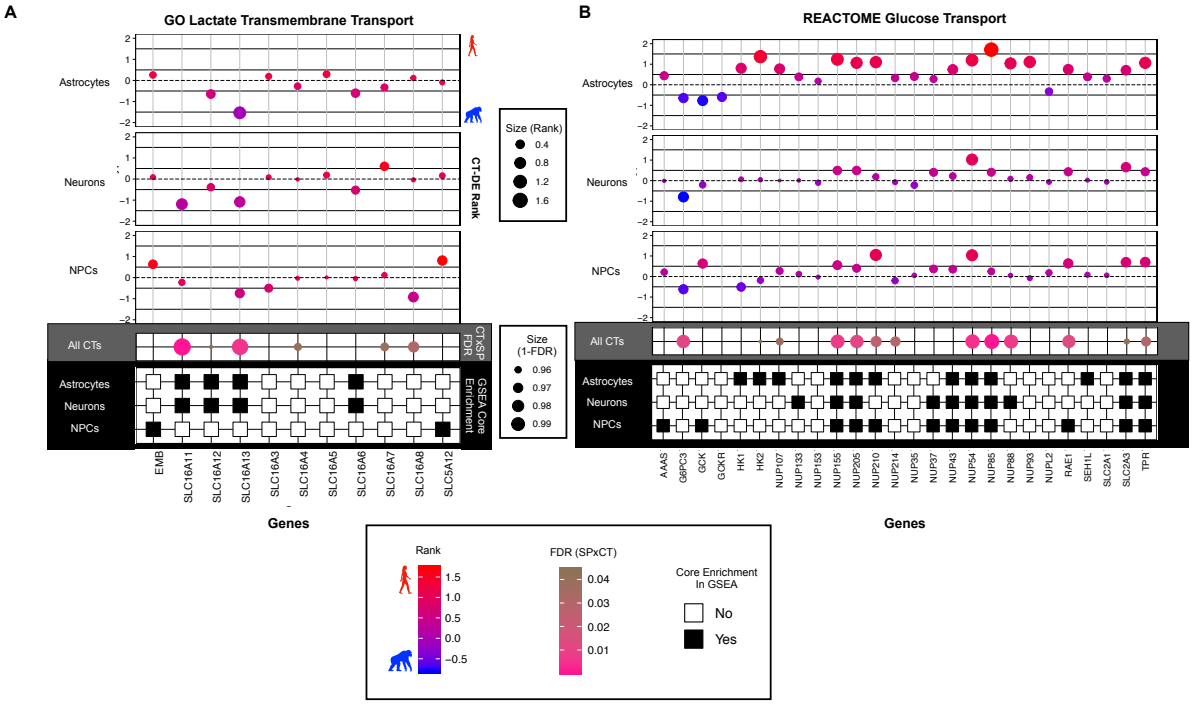

A
GO Lactate Transmembrane Transport
Astrocytes
Neurons
NPCs
All CTs
CT-DE Rank
CTxSP FDR
GSEA Core Enrichment
Size (Rank)
0.4
0.8
1.2
1.6
Size (1-FDR)
0.96
0.97
0.98
0.99
EMB
SLC16A11
SLC16A12
SLC16A13
SLC16A3
SLC16A4
SLC16A5
SLC16A6
SLC16A7
SLC16A8
SLC5A12
Genes
B
REACTOME Glucose Transport
AAAS
G6PC3
GCK
GCKR
HK1
HK2
NUP107
NUP133
NUP153
NUP155
NUP205
NUP210
NUP214
NUP35
NUP37
NUP43
NUP54
NUP85
NUP88
NUP93
NUPL2
RAE1
SEH1L
SLC2A1
SLC2A3
TPR
Genes
Rank
1.5
1.0
0.5
0.0
−0.5
FDR (SPxCT)
0.04
0.03
0.02
0.01
Core Enrichment In GSEA
No
Yes

Figure 11 | Neurons and astrocytes exhibit contrasting interspecies differences in lactate and glucose transport. Comparison of genes involved in the A) GO lactate transmembrane transport and B) REACTOME glucose transport gene sets determined as significantly enriched in one species by GSEA analyses. Each panel (white, grey, and black) indicates significance per gene for one of the functional enrichment analyses used (top/white - interspecies DE by cell type (CT-DE); middle/grey – SPxCT ANOVA-like DE, bottom/black – membership in core enrichment genes of leading edge GSEA analysis). (Top panel) Plot of CT-DE rank per each gene (x-axis), where values greater than zero indicate higher expression in human and values less than zero indicate higher expression in chimpanzee. Color spectrum and size also indicate rank (red – higher in human, blue – higher in chimpanzee, larger = higher rank). (Dark grey panel) Plot of FDR for each gene in the ANOVA-like SPxCT DE comparison, where color indicates significance (pink – lower/significant, brown – higher/non-significant) and size also indicates significance (1-FDR, larger = more significant). (Black panel) Plot of whether each gene was part of the core set of genes in GSEA leading edge analysis (black – yes, white – no). See SI Table 3 for DE results per gene and SI Table 6 for GSEA results per gene.

Table 5

Gene ID	Gene Name	dN/dS	logFC		
			NPC	NEU	AST
DCTN6	dynactin subunit 6	1.59	**-1.33**	**-1.43**	**-1.64**
NBPF3	NBPF member 3	1.54	**-1.48**	-1.09	-0.80
CA2	carbonic anhydrase 2	1.53	**3.81**	4.18	0.47
CDCA5	cell division cycle associated 5	1.52	-0.08	-0.34	**1.51**
HHLA3	HERV-H LTR-associating 3	1.45	**5.35**	**4.93**	**3.97**
DBNDD2	dysbindin domain containing 2	1.16	**-4.15**	**-4.54**	**-3.40**
PCYT1A	phosphate cytidylyltransferase 1, choline, alpha	1.14	**1.29**	**2.16**	**1.56**
HRC	histidine rich calcium binding protein	1.13	**2.12**	*NA*	2.23
LIN9	lin-9 DREAM MuvB core complex component	1.11	**1.32**	1.04	**1.81**
IMPG1	interphotoreceptor matrix proteoglycan 1	1.08	**-4.39**	-2.07	-2.32
EME1	essential meiotic structure-specific endonuclease 1	1.06	-0.75	**-1.47**	0.30
TMEM140	transmembrane protein 140	1.03	**1.88**	1.76	1.48
CD24	CD24 molecule	1.01	**2.66**	**3.15**	2.64

Table 5 | Interspecies differentially expressed genes with evidence of coding selection. In order to determine if genes exhibiting significant interspecies differential expression also had evidence of positive selection in their coding sequences, we used nonsynonymous (dN) and synonymous (dS) nucleotide changes per gene for all genes expressed in iPSC-derived neural cells. A rate of change was calculated for each gene (dN/dS), where a dN/dS > 1 is indicative of positive selection. logFC from CT-DE comparisons are listed for each gene; color indicates degree of logFC and species with higher expression (red, human; blue, chimpanzee). Bolded logFC values indicate they are significantly DE between species in that cell type.

Figure 12

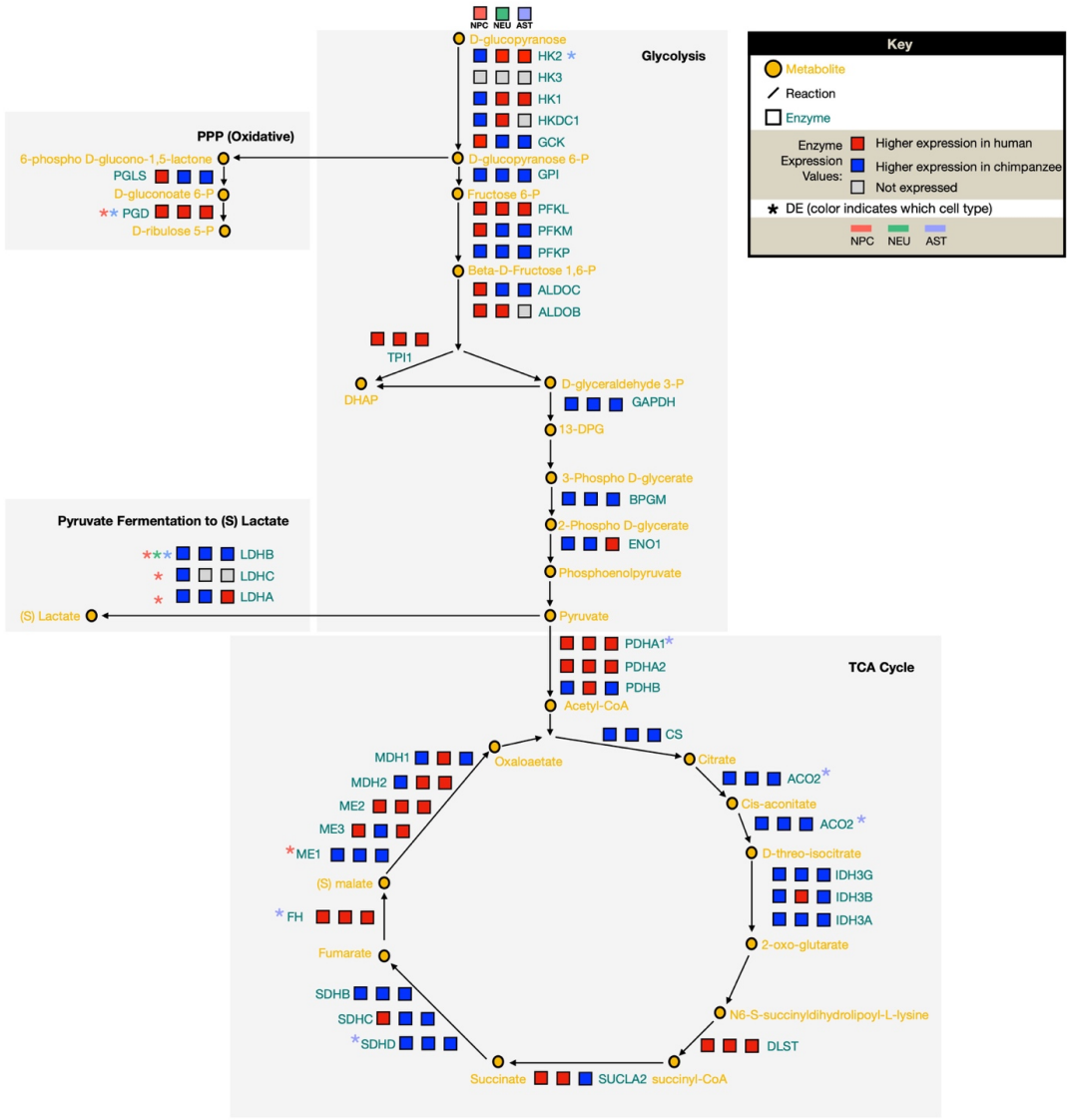
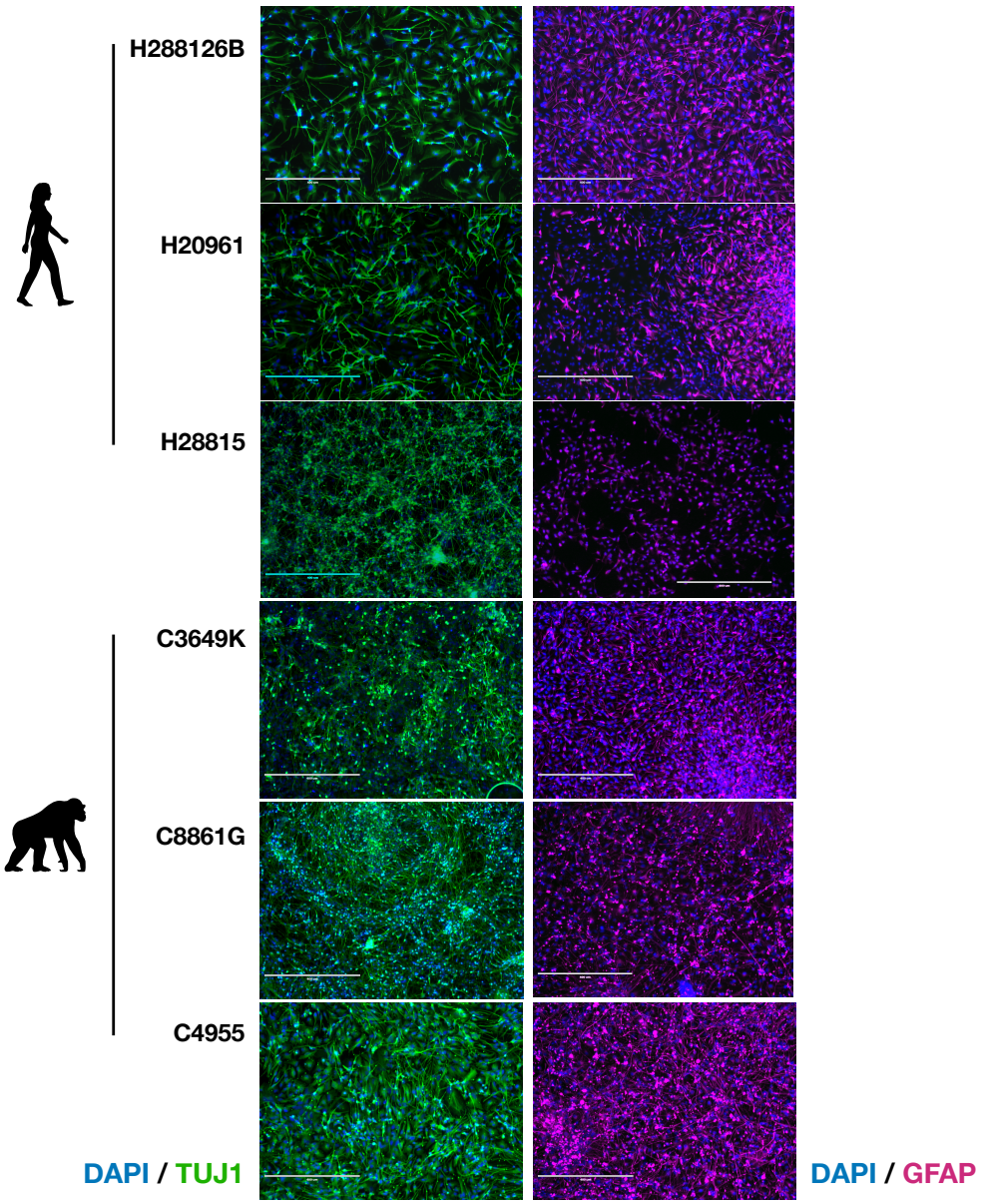

NPC NEU AST
Glycolysis
D-glucopyranose
HK2 *
HK3
HK1
HKDC1
GCK
D-glucopyranose 6-P
GPI
Fructose 6-P
PFKL
PFKM
PFKP
Beta-D-Fructose 1,6-P
ALDOC
ALDOB
TPI1
DHAP
D-glyceraldehyde 3-P
GAPDH
13-DPG
3-Phospho D-glycerate
BPGM
2-Phospho D-glycerate
ENO1
Phosphoenolpyruvate
Pyruvate
PPP (Oxidative)
6-phospho D-glucono-1,5-lactone
PGLS
D-gluconoate 6-P
** PGD
D-ribulose 5-P
Pyruvate Fermentation to (S) Lactate
*** LDHB
* LDHC
* LDHA
(S) Lactate
PDHA1*
PDHA2
PDHB
Acetyl-CoA
CS
TCA Cycle
Oxaloacetate
Citrate
ACO2*
Cis-aconitate
ACO2 *
D-threo-isocitrate
IDH3G
IDH3B
IDH3A
2-oxo-glutarate
N6-S-succinyldihydrolipoyl-L-lysine
DLST
Succinate
SUCLA2 succinyl-CoA
*SDHD
SDHC
SDHB
Fumarate
* FH
(S) malate
*ME1
ME3
ME2
MDH2
MDH1
Key
Metabolite
Reaction
Enzyme
Enzyme Expression Values:
Higher expression in human
Higher expression in chimpanzee
Not expressed
★ DE (color indicates which cell type)
NPC NEU AST

Figure 12 | Divergence in pyruvate utilization between species' astrocytes. We constructed a focal set of aerobic glycolysis signaling pathways in order to contextualize our DE results in the framework of a network signaling. A diagram of the major pathways involved in aerobic glycolysis (glycolysis, pentose phosphate pathway (PPP), lactate conversion from pyruvate, and TCA cycle). For each enzyme in the pathway, three blocks indicate expression of this enzyme in each cell type – left to right: NPCs, neurons, astrocytes. Color indicates level of expression (higher in human (red), higher in chimpanzee (blue), not expressed in this cell type (grey)).

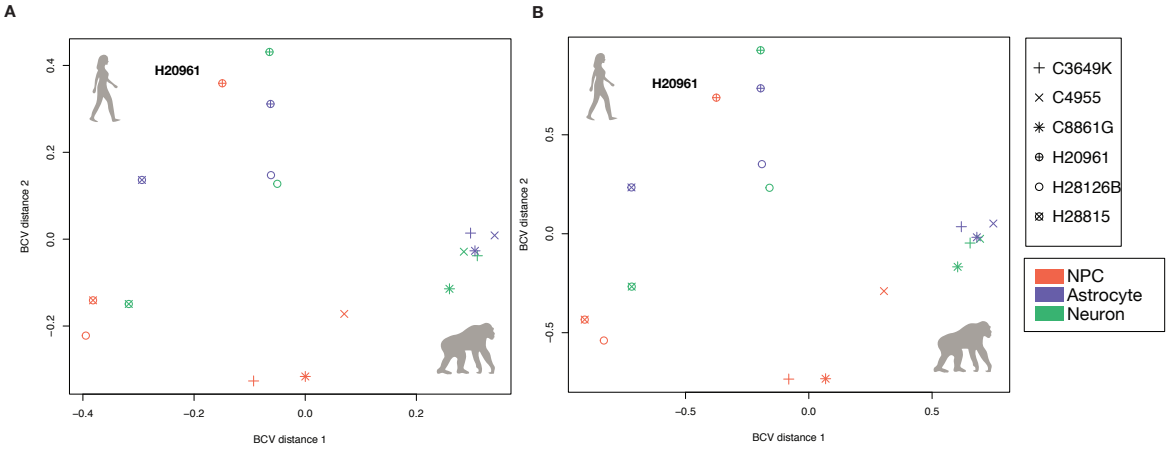
3.9 Supplemental Figures and Tables

SI Figure 13



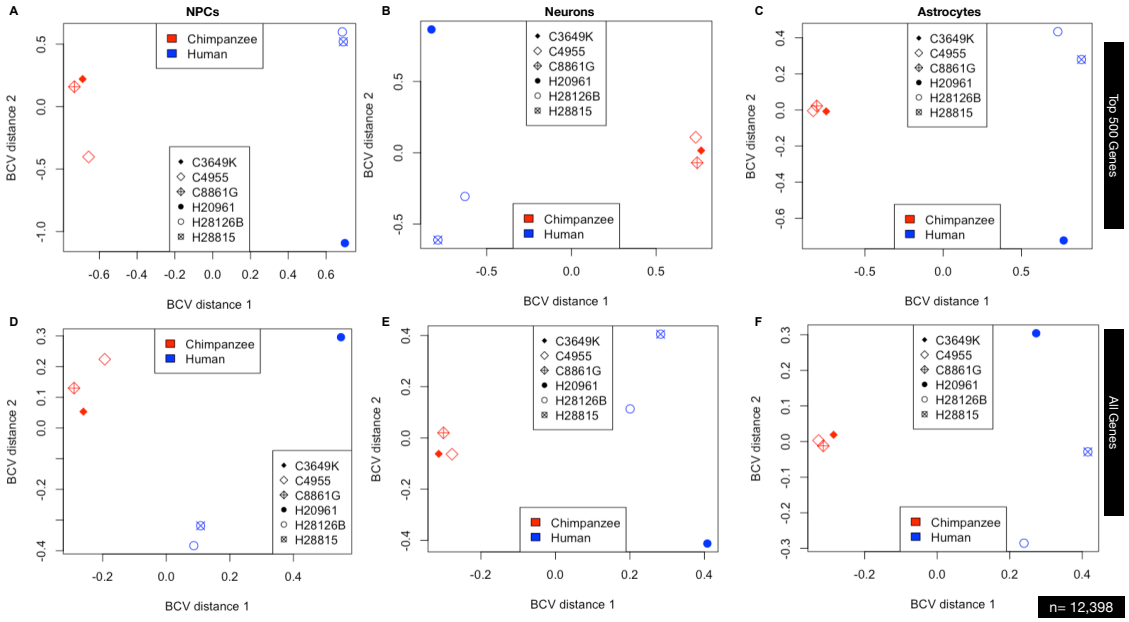
SI Figure 13 | Differentiation and maturation of a human and chimpanzee iPSC lines into neural cell types. Immunofluorescent validation of matched-by-cell-line iPSC-derived neurons and astrocytes. Left column is images of cells immunofluorescently labeled for neuron-specific class III β-tubulin (TUJ1; Neuromics), and the right column is images of iPSC-derived astrocytes immunofluorescently labeled for GFAP (Sigma Aldrich), according to manufacturer's suggestions. All cells for each cell type were harvested at similar timepoints: neurons at passage 3-4 and mature astrocytes at passage 5-6 post-differentiation from NPCs (SI Table 1).

SI Figure 14


A
B
H20961
H20961
BCV distance 1
BCV distance 2
BCV distance 1
BCV distance 2
C3649K
C4955
C8861G
H20961
H28126B
H28815
NPC
Astrocyte
Neuron

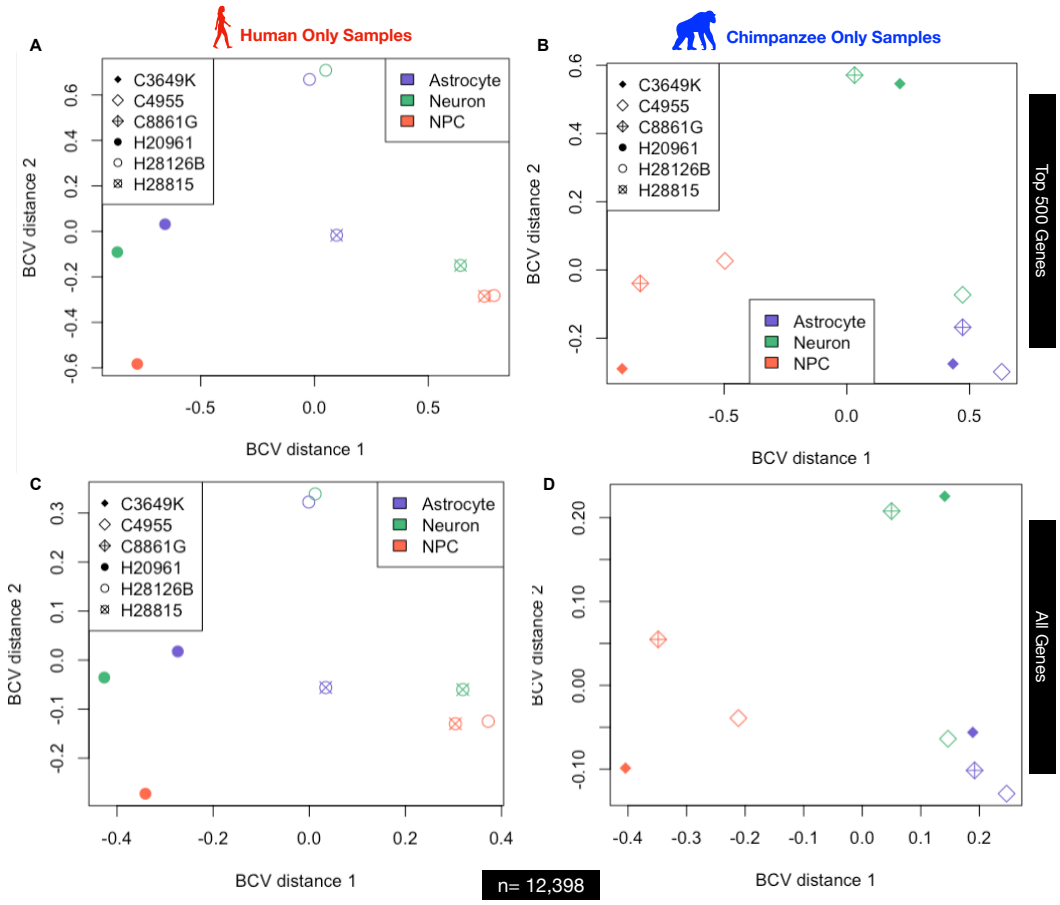
SI Figure 14 | MDS plots of all iPSC-derived samples with shape indicating cell line. The same MDS plots for A) all genes expressed and B) the top 500 more differentially expressed genes as in main Figure 1C where shape indicates individual cell lines.

SI Figure 15


A
NPCs
B
Neurons
C
Astrocytes
D
E
F
Chimpanzee
Human
C3649K
C4955
C8861G
H20961
H28126B
H28815
BCV distance 1
BCV distance 2
Top 500 Genes
All Genes
n= 12,398

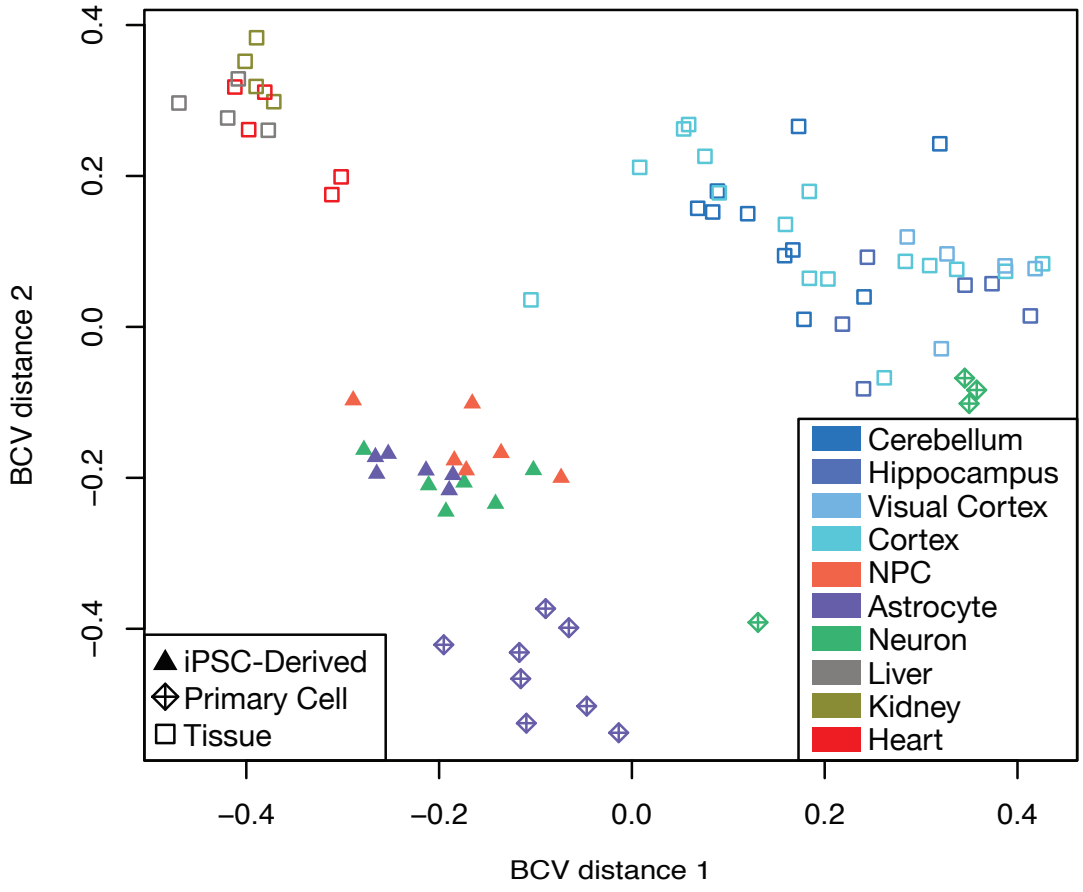
SI Figure 15 | MDS plots of individual cell types (A & D – NPCs, B & E – neurons, C & F – astrocytes). Plots A-C are for the top 500 most differentially expressed genes while plots D-F are for all genes expressed.

SI Figure 16


A
Human Only Samples
B
Chimpanzee Only Samples
C3649K
C4955
C8861G
H20961
H28126B
H28815
Astrocyte
Neuron
NPC
BCV distance 1
BCV distance 2
Top 500 Genes
C
D
All Genes
n= 12,398

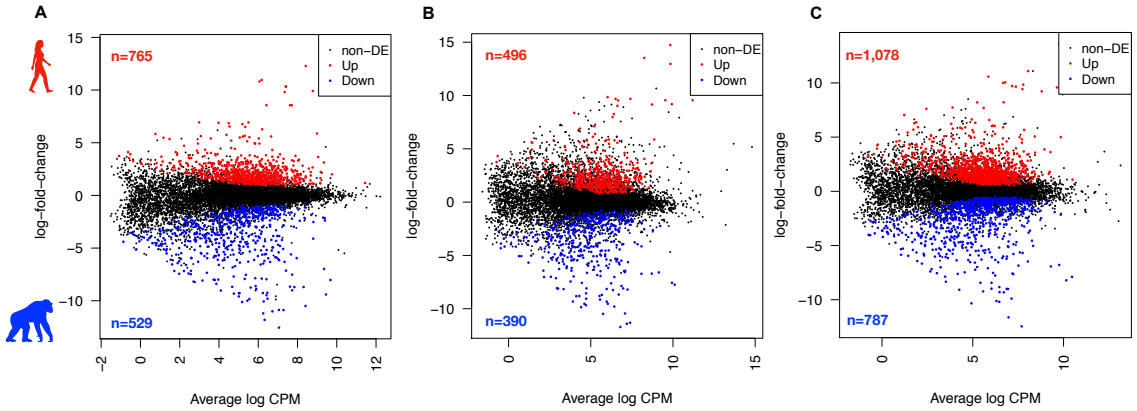
SI Figure 16 | MDS plots of all A & C) human and B & D) chimpanzee samples by cell type. Plots A & B are for the top 500 most differentially expressed genes while plots C & D are for all genes expressed.

SI Figure 17


BCV distance 2
0.4
0.2
0.0
−0.2
−0.4
BCV distance 1
−0.4
−0.2
0.0
0.2
0.4
iPSC-Derived
Primary Cell
Tissue
Cerebellum
Hippocampus
Visual Cortex
Cortex
NPC
Astrocyte
Neuron
Liver
Kidney
Heart

SI Figure 17 | Human and chimpanzee iPSC-derived neural cells resemble primary neural cell types and tissue regions more than non-neuronal tissues. A PCoA of the iPSC-derived neural cells in comparison to whole-tissue RNA-Seq from four brain regions (cerebellum, hippocampus, prefrontal cortex, and visual cortex) from human and chimpanzee (3 individuals per species) (Babbitt et al , *in prep.*), brain and non-neuronal tissue from human and chimpanzee from [45], as well as that from primary neurons and astrocytes [186-188] obtained from the Gene Expression Omnibus (GEO) [185] Short Read Archive (SRA) (GEO accession numbers GSE30352, GSE73721, GSM2071331, GSM2071332, and GSM2071418). Label shape indicates the sample source tissue or cell type. Color refers to cell type or brain region.

SI Figure 18



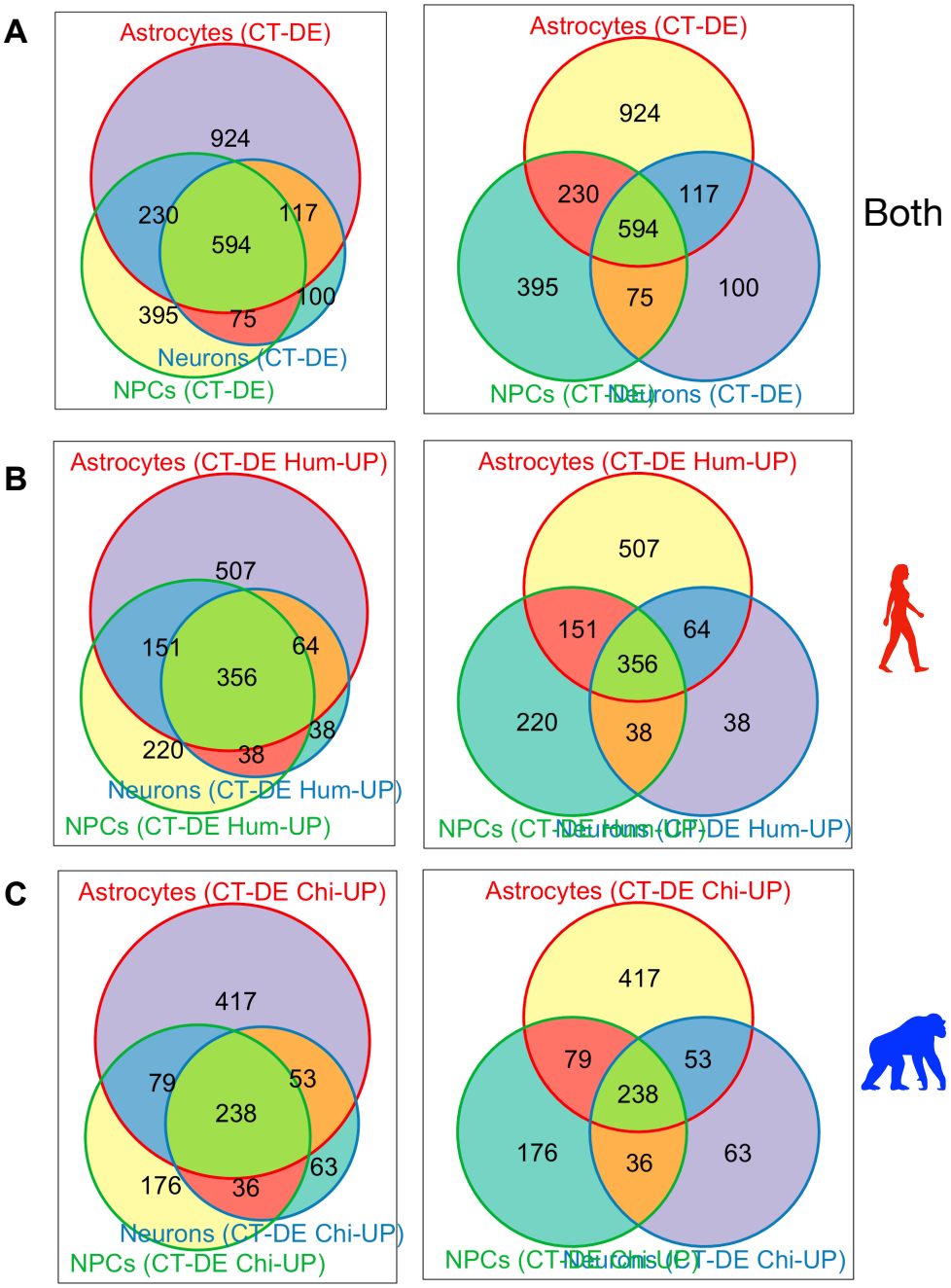
DE Total: 1,294
Total genes expr. in all NPCs: 12,481
Genes after filtering CPM > 1: 11,772
% DE:
Ratio (H/C): 1.45

DE Total: 886
Total genes expr. in all Neurons: 12,451
Genes after filtering CPM > 1: 12,046
% DE:
Ratio (H/C): 1.27

DE Total: 1,865
Total genes expr. in all Astrocytes: 12,302
Genes after filtering CPM > 1: 11,854
% DE:
Ratio (H/C): 1.34

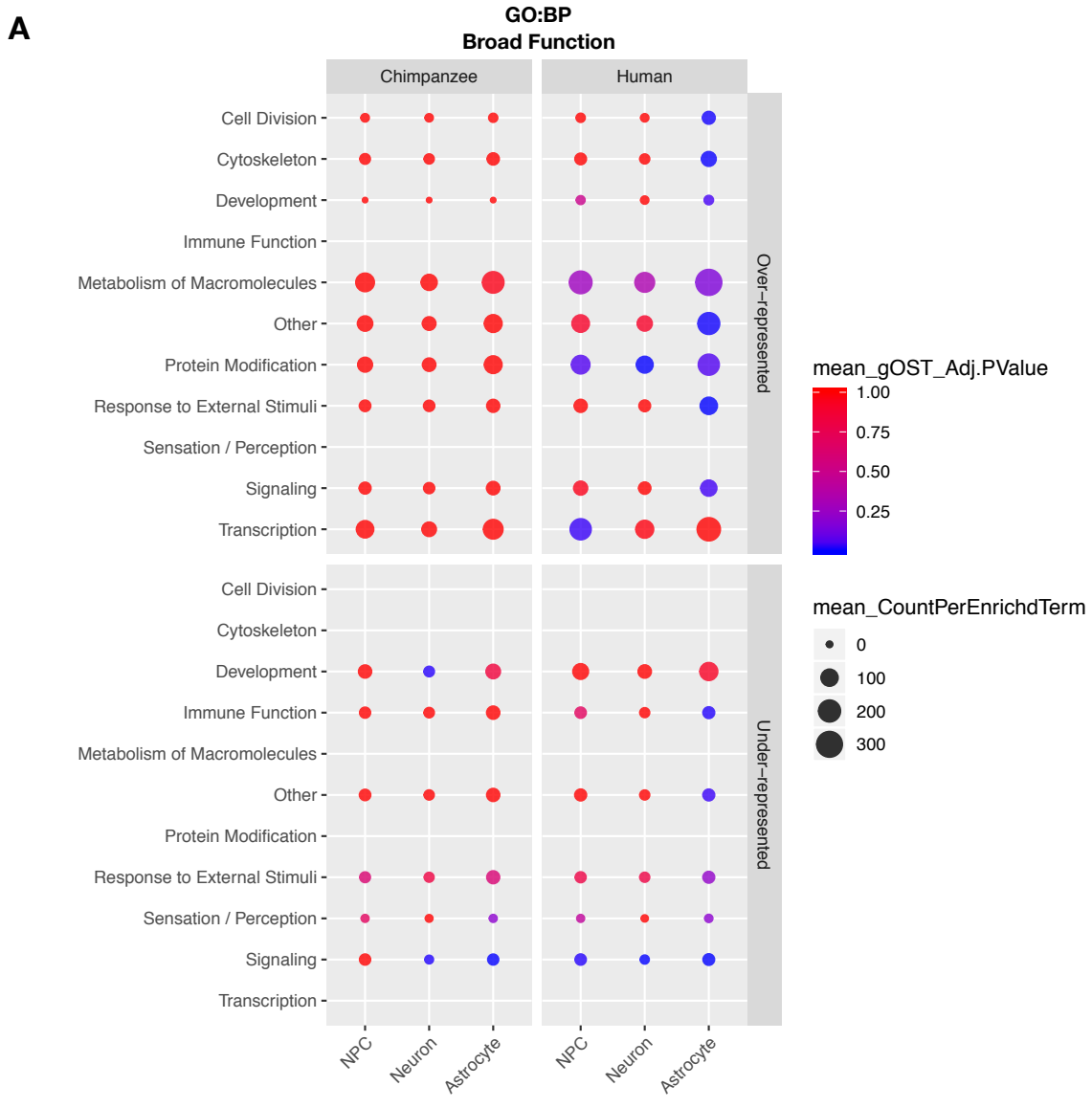
SI Figure 18 | Distribution of differentially expressed genes between species for each cell type. MA plots of expression of all genes (average log CPM, x-axis) by their relative log fold-change (y-axis) from TopTags tables of pairwise, interspecies CT-DE comparisons made in edgeR for A) NPCs, B) neurons, and C) astrocytes. Color indicates differential expression status: black – non-DE, red – DE with higher expression in human, blue – DE with higher expression in chimpanzee). Number of genes identified as differentially expressed with higher expression in human and indicated in red text; for chimpanzee, in blue text.

SI Figure 19


A
Astrocytes (CT-DE)
924
230
117
594
100
395
75
Neurons (CT-DE)
NPCs (CT-DE)
Astrocytes (CT-DE)
924
230
117
594
395
75
100
NPCs (CT-DE)
Neurons (CT-DE)
Both
B
Astrocytes (CT-DE Hum-UP)
507
151
64
356
38
220
38
Neurons (CT-DE Hum-UP)
NPCs (CT-DE Hum-UP)
Astrocytes (CT-DE Hum-UP)
507
151
64
356
220
38
38
NPCs (CT-DE Hum-UP)
Neurons (CT-DE Hum-UP)
C
Astrocytes (CT-DE Chi-UP)
417
79
53
238
63
176
36
Neurons (CT-DE Chi-UP)
NPCs (CT-DE Chi-UP)
Astrocytes (CT-DE Chi-UP)
417
79
53
238
176
36
63
NPCs (CT-DE Chi-UP)
Neurons (CT-DE Chi-UP)

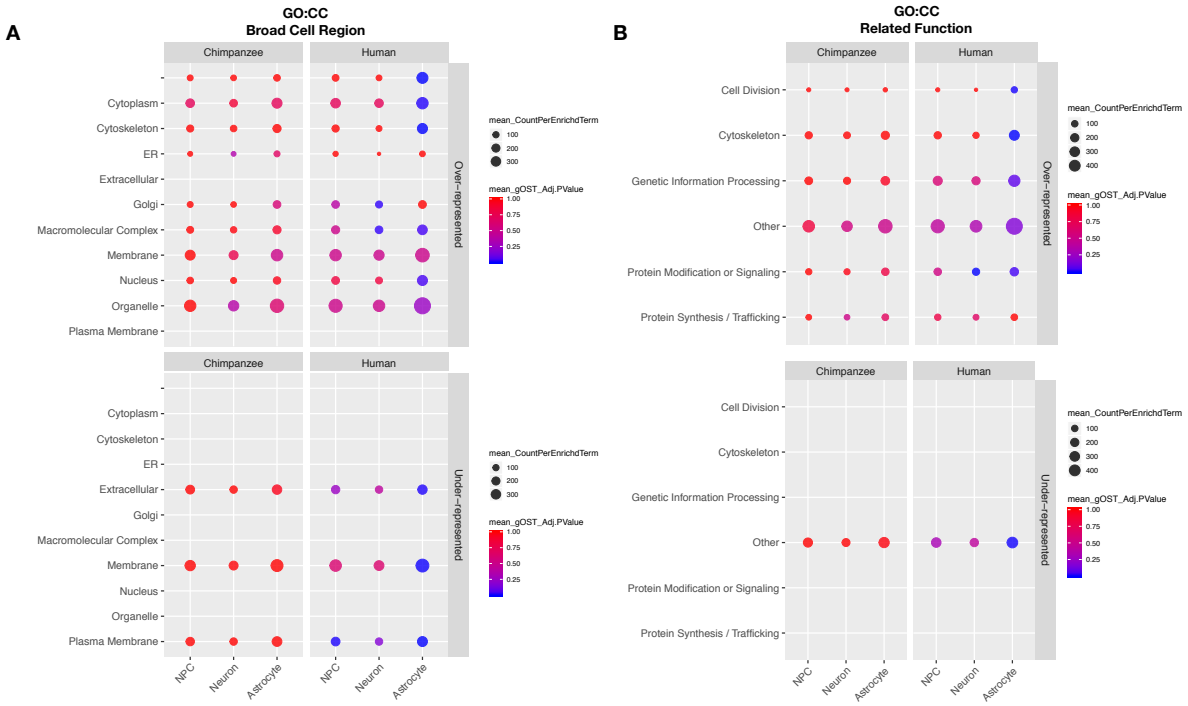
SI Figure 19 | Overlap in interspecies CT-DE genes. Weighted (left column) and unweighted (right column) Venn diagrams made using R package Vennerable of overlap in genes per cell type exhibiting differential expression between species. Venn diagrams are for all CT-DE genes across all CT's (A) and DE genes across all CT-DE comparisons with higher expression in human (B) or chimpanzee (C).

SI Figure 20


A
GO:BP
Broad Function
Chimpanzee
Human
Cell Division
Cytoskeleton
Development
Immune Function
Metabolism of Macromolecules
Other
Protein Modification
Response to External Stimuli
Sensation / Perception
Signaling
Transcription
Over−represented
Under−represented
NPC
Neuron
Astrocyte
mean_gOST_Adj.PValue
1.00
0.75
0.50
0.25
mean_CountPerEnrichdTerm
0
100
200
300

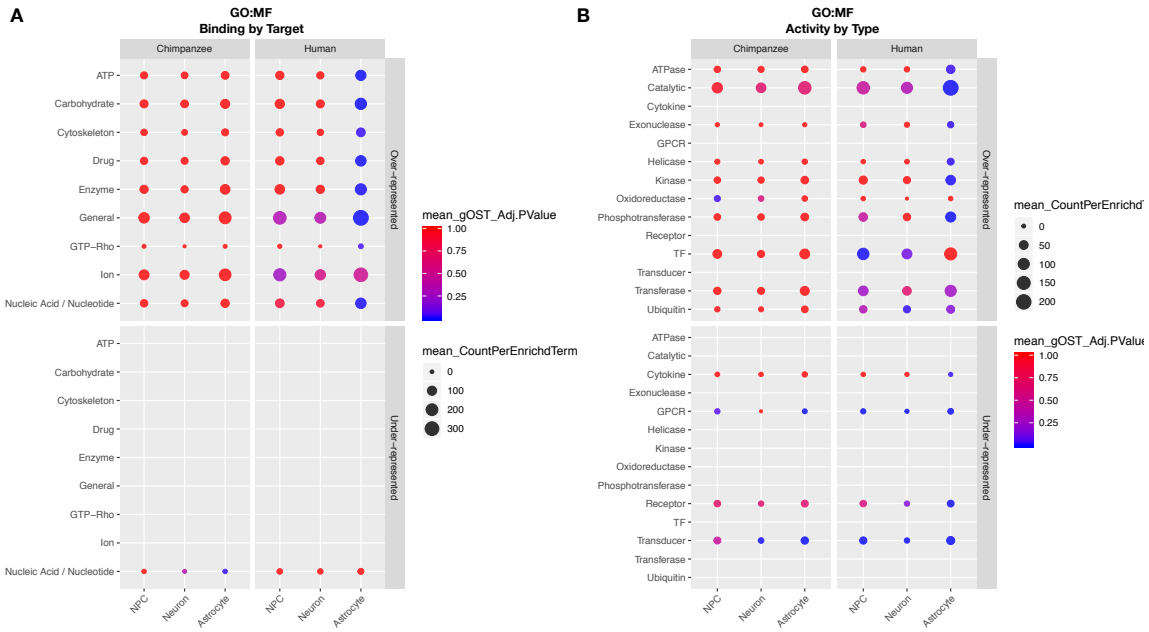
SI Figure 20 | GO Biological Process (BP) enrichments. Plots of significantly over-represented (top panels) and under-represented (bottom panels) categories of GO BP terms determined by categorical enrichment analyses in genes with higher expression in chimpanzee (left panel) and human (right panel) for each cell type (x-axis). The categories (y-axis) represent groupings of multiple GO BP terms grouped by their general function. Size indicates the mean count and color indicates the mean adjusted enrichment p-value for all terms in that category.

SI Figure 21


A
GO:CC
Broad Cell Region
Chimpanzee
Human
Cytoplasm
Cytoskeleton
ER
Extracellular
Golgi
Macromolecular Complex
Membrane
Nucleus
Organelle
Plasma Membrane
Over−represented
Under−represented
mean_CountPerEnrichdTerm
100
200
300
mean_gOST_Adj.PValue
1.00
0.75
0.50
0.25
NPC
Neuron
Astrocyte
B
GO:CC
Related Function
Cell Division
Cytoskeleton
Genetic Information Processing
Other
Protein Modification or Signaling
Protein Synthesis / Trafficking
400

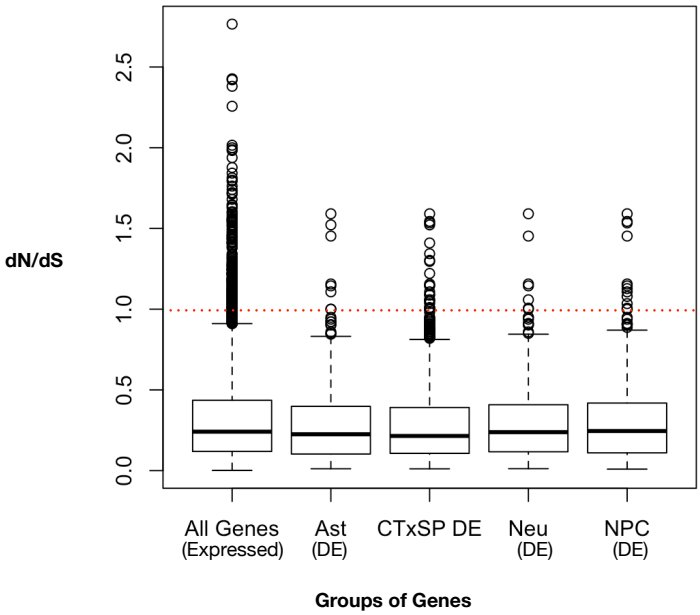
SI Figure 21 | GO Cellular Component (CC) enrichments. Plots of significantly over-represented (top panels) and under-represented (bottom panels) categories of GO CC terms determined by categorical enrichment analyses in genes with higher expression in chimpanzee (left panel) and human (right panel) for each cell type (x-axis). The categories (y-axis) represent groupings of multiple GO CC terms grouped by their A) by their broad cell region and B) functions related to the cellular components enriched. Size indicates the mean count and color indicates the mean adjusted enrichment p-value for all terms in that category.

SI Figure 22


A
GO:MF
Binding by Target
Chimpanzee
Human
ATP
Carbohydrate
Cytoskeleton
Drug
Enzyme
General
GTP–Rho
Ion
Nucleic Acid / Nucleotide
Over−represented
Under−represented
NPC
Neuron
Astrocyte
mean_gOST_Adj.PValue
1.00
0.75
0.50
0.25
mean_CountPerEnrichdTerm
0
100
200
300
B
GO:MF
Activity by Type
Chimpanzee
Human
ATPase
Catalytic
Cytokine
Exonuclease
GPCR
Helicase
Kinase
Oxidoreductase
Phosphotransferase
Receptor
TF
Transducer
Transferase
Ubiquitin
Over−represented
Under−represented
NPC
Neuron
Astrocyte
mean_CountPerEnrichd
0
50
100
150
200
mean_gOST_Adj.PValue
1.00
0.75
0.50
0.25

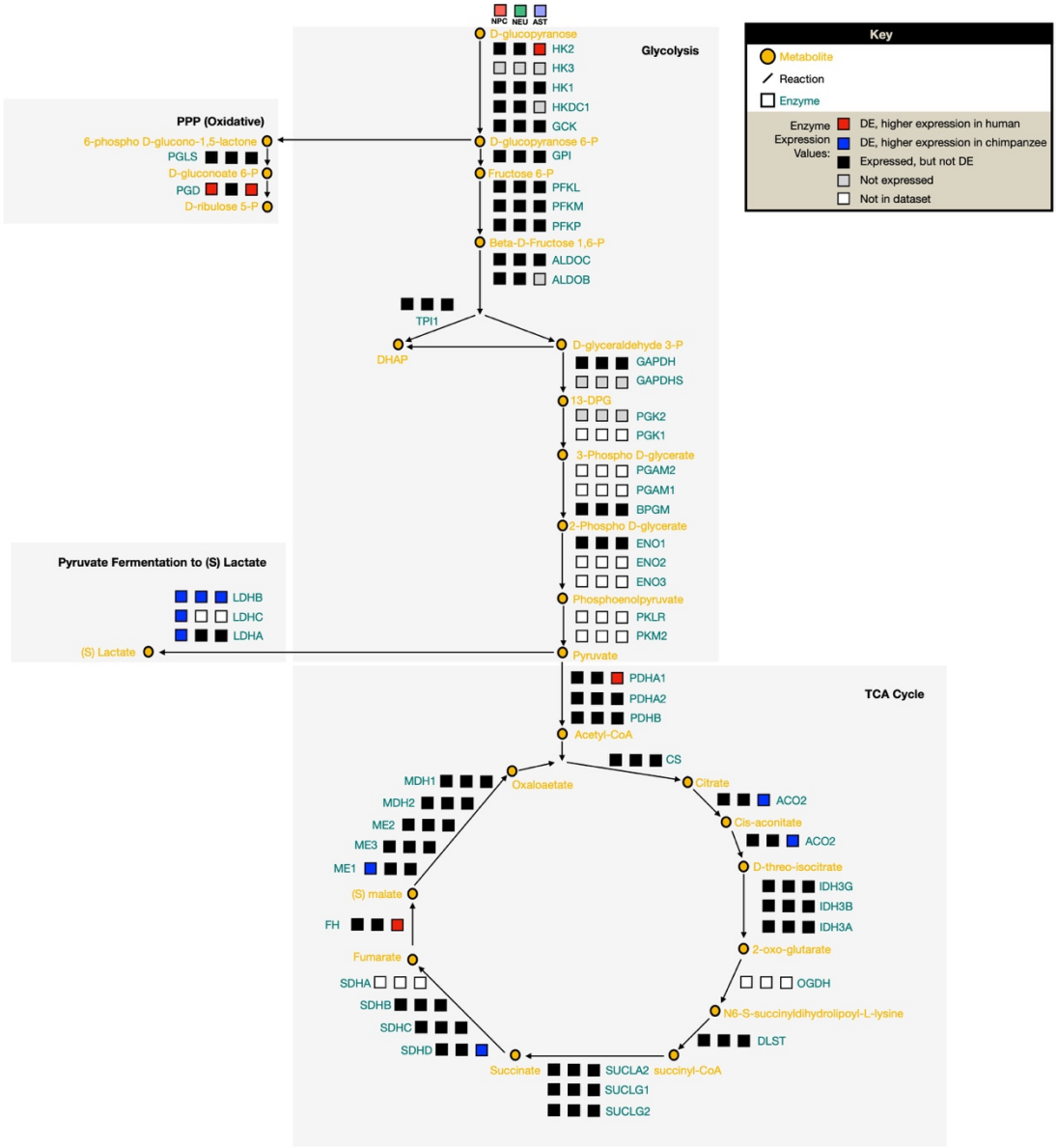
SI Figure 22 | GO Molecular Function (MF) enrichments. Plots of significantly over-represented (top panels) and under-represented (bottom panels) categories of GO MF terms determined by categorical enrichment analyses in genes with higher expression in chimpanzee (left panel) and human (right panel) for each cell type (x-axis). The categories (y-axis) represent groupings of multiple GO MF terms grouped by their A) binding activity for particular substrates and B) specificity types of molecular activity. Size indicates the mean count and color indicates the mean adjusted enrichment p-value for all terms in that category.

SI Figure 23


dN/dS
0.0
0.5
1.0
1.5
2.0
2.5
All Genes (Expressed)
Ast (DE)
CTxSP DE
Neu (DE)
NPC (DE)
Groups of Genes

SI Figure 23 | Positive selection in genes expressed in iPSC-derived neural cells. In order to determine if genes exhibiting significant interspecies differential expression also had evidence of positive selection in their coding sequences, we used nonsynonymous (dN) and synonymous (dS) nucleotide changes per gene for all genes expressed in iPSC-derived neural cells. A rate of change was calculated for each gene (dN/dS), where a dN/dS > 1 is indicative of positive selection.

SI Figure 24


NPC NEU AST
Glycolysis
D-glucopyranose
HK2
HK3
HK1
HKDC1
GCK
D-glucopyranose 6-P
GPI
Fructose 6-P
PFKL
PFKM
PFKP
Beta-D-Fructose 1,6-P
ALDOC
ALDOB
TPI1
DHAP
D-glyceraldehyde 3-P
GAPDH
GAPDHS
13-DPG
PGK2
PGK1
3-Phospho D-glycerate
PGAM2
PGAM1
BPGM
2-Phospho D-glycerate
ENO1
ENO2
ENO3
Phosphoenolpyruvate
PKLR
PKM2
Pyruvate
PPP (Oxidative)
6-phospho D-glucono-1,5-lactone
PGLS
D-gluconoate 6-P
PGD
D-ribulose 5-P
Pyruvate Fermentation to (S) Lactate
LDHB
LDHC
LDHA
(S) Lactate
PDHA1
PDHA2
PDHB
Acetyl-CoA
TCA Cycle
CS
Citrate
ACO2
Cis-aconitate
ACO2
D-threo-isocitrate
IDH3G
IDH3B
IDH3A
2-oxo-glutarate
OGDH
N6-S-succinyldihydrolipoyl-L-lysine
DLST
succinyl-CoA
SUCLA2
SUCLG1
SUCLG2
Succinate
SDHA
SDHB
SDHC
SDHD
Fumarate
FH
(S) malate
MDH1
MDH2
ME2
ME3
ME1
Oxaloacetate
Key
Metabolite
Reaction
Enzyme
Enzyme Expression Values:
DE, higher expression in human
DE, higher expression in chimpanzee
Expressed, but not DE
Not expressed
Not in dataset

SI Figure 24 | Full expression network of sub-pathways in aerobic glycolysis. We constructed a focal set of aerobic glycolysis signaling pathways in order to contextualize our DE results in the framework of a network signaling. A diagram of the major pathways involved in aerobic glycolysis (glycolysis, pentose phosphate pathway (PPP), lactate conversion from pyruvate, and TCA cycle). For each enzyme in the pathway, three blocks indicate expression of this enzyme in each cell type – left to right: NPCs, neurons, astrocytes. Color indicates level of expression (DE and higher in human (red), DE and higher in chimpanzee (blue), not expressed in this cell type (grey), expressed but not DE (black)).

CONCLUSIONS

My dissertation work has significantly increased the taxonomic, tissue, and cell type breadth of primate transcriptome data. It includes the first in-depth investigation into V1 expression changes across such a phenotypically and phylogenetically distinct set of primates and demonstrates that alterations in metabolic signaling play an important role in evolved differences in V1 gene expression. These differences in V1 expression likely overlap with that of other brain regions, given that it is a cortical region that as a whole has experienced significant evolutionary change over primate evolution, particularly in the human lineage. However, given the significant differences observed even between the relatively recently diverged human and chimpanzee, some of these metabolic and neuron-specific signaling changes may be uniquely important to V1. Future studies with more sophisticated analytical approaches may be able to determine this.

Despite the known importance of vision in primate evolution, my first chapter represents one of very few investigations of the V1 region at the level of gene expression. V1 has historically been well-studied at the tissue level, particularly in rhesus macaques (*Macaca mulatta*), but any previous investigations of V1 at the expression level have been limited in species number and variability (to human, chimpanzee, and macaque) and have largely been more focused on other brain regions. These analyses, though nuanced and not without important caveats, also represent one of few studies attempting to link gene expression changes in a visual-stimuli sensing brain region with that of visually-relevant, putatively adaptive primate phenotypes. Our finding of significant expression differences in V1 associated with differences in a proxy for sociality and the inclusion of genes of known function in sociality (*OXTR* and *DRD4*) are particularly intriguing. This work also demonstrates that despite significant conservation in both DNA nucleotide sequence and visual processing systems between humans

and chimpanzees, gene expression in V1 is substantially more different between these species than other species.

There are a number of ethical and practical hurdles to investigating primate tissue and cell samples. Additional novelty of my dissertation research has been taking an interdisciplinary approach to investigating classic questions in biological anthropology – such as interspecies differences in organism-level traits such as disease and brain metabolic demands – using contemporary molecular and cellular *in vitro* techniques. The utility of relatively non-invasive skin biopsies to obtain fibroblasts from humans and chimpanzees has been key to my investigation of cell adhesion and migration in fibroblasts as well as neural-cell type specific expression changes, given that the iPSCs were generated from fibroblast lines. Cellular level and cell-type specific studies of evolved changes between species are fairly rare, and the contributions of my dissertation to this niche field are significant. They serve to further elucidate the proximate molecular mechanisms putatively influencing ultimate evolutionary trajectories of human-specific evolution.

The fairly new field of evolutionary medicine aims to better understand human disease by taking a comparative biology approach using the "natural experiment" of primate evolution. However, studying species-level differences in complex disease phenotypes is inherently difficult, for a number of practical (e.g. rarity of disease samples or individuals) or ethical (e.g. laboratory experiments on human and non-human primates) reasons. My investigation of basic biological differences in cell adhesion and migration represent a way to test hypotheses regarding the molecular and cellular mechanisms that contribute to these species-level differences. I determined that previously determined significant interspecies differences in focal adhesion complexes in fibroblasts correlate with significant changes in in cell adhesion and migration. In conjunction with Jason Pizzollo and colleagues, I also determined that these changes correlate not only with altered focal adhesion gene expression during a wound healing response, but by consistent increased expression of focal adhesion and cellular adhesion

molecule genes by chimpanzee fibroblasts. Not only do these results inform the field of evolutionary biology, but also provide putative pathway and gene level targets of human-specific altered expression that may play important roles in the development of human epithelial neoplasms.

Finally, our use of modern iPSC and cell differentiation technology has allowed us to study specific neural cell types from both humans and chimpanzees. Importantly, we are the only lab known to date to investigate astrocytes from chimpanzees. Astrocytes are a historically understudied cell type in the brain, and recently, have been determined to have important functions in synapse development and neural signaling, in addition to critical metabolic support, and aberrations of these cell-type specific functions have been implicated in human neurological disease. Human-specific differences in astrocyte function could further elucidate mechanisms of human-specific disease. My results provide valuable insights into astrocyte-mediated differences in metabolic brain function as a proximate mechanism by which the ultimate evolutionary trajectory of human brain evolution may have occurred. Future studies investigating the overlap in interspecies DE genes of astrocytes with signatures of molecular selection in *cis*-regulatory regions can better determine the adaptive significance of this altered expression. This could include investigations of epigenetic changes in chromatin accessibility (e.g. ATAC-Seq).

The finding of species differences in pyruvate utilization across all neural cell types, with humans exhibiting a greater capacity to produce acetyl-CoA while chimpanzees exhibit higher expression of enzymes conferring conversion of pyruvate into lactate, could be further investigated in a number of ways. RNA-Seq of co-cultured neurons and astrocytes would be an important additional dataset for better understanding interspecies differences in the astrocyte-neuron lactate shuttle. Metabolic profiling of each cell type using high-throughput cellular assays (e.g. Agilent's Seahorse glycolytic rate) could provide cellular-level evidence for interspecies differences in cellular respiration. Finally, increasing the number of species

considered - particularly of iPSC-derived astrocytes – to include rhesus macaque, common marmoset, mice, and rats would allow us to better understand not only the evolutionary pattern of altered aerobic glycolysis gene expression but also could elucidate how appropriate each of these biomedical model organisms are for studying human-specific neural function.

BIBLIOGRAPHY

1. Preuss, T.M., *Human brain evolution: From gene discovery to phenotype discovery.* Proceedings of the National Academy of Sciences of the United States of America, 2012. **109**: p. 10709-10716.

2. Goldberg, A., et al., *Adaptive evolution of cytochrome c oxidase subunit VIII in anthropoid primates.* Proceedings of the National Academy of Sciences of the United States of America, 2003. **100**(10): p. 5873-8.

3. Grossman, L.I., et al., *Molecular evolution of aerobic energy metabolism in primates.* Mol Phylogenet Evol, 2001. **18**(1): p. 26-36.

4. Grossman, L.I., et al., *Accelerated evolution of the electron transport chain in anthropoid primates.* Trends Genet, 2004. **20**(11): p. 578-85.

5. Huttemann, M., et al., *Regulation of mitochondrial respiration and apoptosis through cell signaling: cytochrome c oxidase and cytochrome c in ischemia/reperfusion injury and inflammation.* Biochim Biophys Acta, 2012. **1817**(4): p. 598-609.

6. Uddin, M., et al., *Molecular evolution of the cytochrome c oxidase subunit 5A gene in primates.* BMC Evol Biol, 2008. **8**: p. 8.

7. Mollon, J.D., *"Tho'she kneel'd in that place where they grew…" The uses and origins of primate colour vision.* Journal of Experimental Biology, 1989. **146**(1): p. 21-38.

8. Pessoa, D.M.A., et al., *The adaptive value of primate color vision for predator detection.* American Journal of Primatology, 2014. **76**(8): p. 721-729.

9. Dominy, N.J. and P.W. Lucas, *Ecological importance of trichromatic vision to primates.* Nature, 2001. **410**(6826): p. 363.

10. Jacobs, G.H., *Primate color vision: a comparative perspective.* Visual Neuroscience, 2008. **25**(5-6): p. 619-633.

11. Martin, R.D. and C.F. Ross, *The evolutionary and ecological context of primate vision.* The primate visual system: A comparative approach, 2005: p. 1-36.

12. Hall, M.I., J.M. Kamilar, and E.C. Kirk, *Eye shape and the nocturnal bottleneck of mammals.* Proceedings of the Royal Society B: Biological Sciences, 2012: p. rspb20122258.

13. Heesy, C.P. and C.F. Ross, *Evolution of activity patterns and chromatic vision in primates: morphometrics, genetics and cladistics.* Journal of Human Evolution, 2001. **40**(2): p. 111-149.

14. Melin, A.D., et al., *Euarchontan opsin variation brings new focus to primate origins.* Molecular Biology and Evolution, 2016. **33**(4): p. 1029-1041.

15. Srinivasan, S., C.N. Carlo, and C.F. Stevens, *Predicting visual acuity from the structure of visual cortex.* Proceedings of the National Academy of Sciences of the United States of America, 2015. **112**(25): p. 7815-7820.

16. Wheeler, B.C., B.J. Bradley, and J.M. Kamilar, *Predictors of orbital convergence in primates: A test of the snake detection hypothesis of primate evolution.* Journal of Human Evolution, 2011. **61**(3): p. 233-242.

17. Joffe, B., et al., *Diurnality and nocturnality in primates: an analysis from the rod photoreceptor nuclei perspective.* Evolutionary Biology, 2014. **41**(1): p. 1-11.

18. Baker, C.I., *Visual processing in the primate brain.* Handbook of Psychology, Second Edition, 2012. **3**.

19. Walls, G., *The Vertebrate Eye and Its Adaptive Radiation. Facsimile edition*. 1942, Hafner Publishing Co., New York.

20. Kay, R.F. and E.C. Kirk, *Osteological evidence for the evolution of activity pattern and visual acuity in primates.* American Journal of Physical Anthropology: The Official Publication of the American Association of Physical Anthropologists, 2000. **113**(2): p. 235-262.

21. Ross, C.F., *Into the light: the origin of Anthropoidea.* Annual Review of Anthropology, 2000. **29**(1): p. 147-194.

22. Ross, C.F. and E.C. Kirk, *Evolution of eye size and shape in primates.* Journal of Human Evolution, 2007. **52**(3): p. 294-313.

23. Kirk, E.C. and R.F. Kay, *The evolution of high visual acuity in the Anthropoidea*, in *Anthropoid Origins*. 2004, Springer. p. 539-602.

24. Kawamura, S., *Color vision diversity and significance in primates inferred from genetic and field studies.* Genes & Genomics, 2016. **38**(9): p. 779-791.

25. Hiramatsu, C., et al., *Experimental evidence that primate trichromacy is well suited for detecting primate social colour signals.* Proceedings of the Royal Society B: Biological Sciences, 2017. **284**(1856): p. 20162458.

26. Surridge, A.K., D. Osorio, and N.I. Mundy, *Evolution and selection of trichromatic vision in primates.* Trends in Ecology & Evolution, 2003. **18**(4): p. 198-205.

27. Osorio, D. and M. Vorobyev, *Colour vision as an adaptation to frugivory in primates.* Proceedings of the Royal Society B: Biological Sciences, 1996. **263**(1370): p. 593-599.

28. Veilleux, C.C., et al., *Group benefit associated with polymorphic trichromacy in a Malagasy primate (Propithecus verreauxi).* Scientific Reports, 2016. **6**: p. 38418.

29. Changizi, M.A., Q. Zhang, and S. Shimojo, *Bare skin, blood and the evolution of primate colour vision.* Biology Letters, 2006. **2**(2): p. 217-221.

30. Fedigan, L.M., et al., *The heterozygote superiority hypothesis for polymorphic color vision is not supported by long-term fitness data from wild neotropical monkeys.* PloS one, 2014. **9**(1): p. e84872.

31. Clifford, C.W. and M. Ibbotson, *Fundamental mechanisms of visual motion detection: models, cells and functions.* Progress in Neurobiology, 2002. **68**(6): p. 409-437.

32. Barton, R.A., *Binocularity and brain evolution in primates.* Proceedings of the National Academy of Sciences of the United States of America, 2004. **101**(27): p. 10113-10115.

33. Preuss, T.M. and G.Q. Coleman, *Human-specific organization of primary visual cortex: alternating compartments of dense Cat-301 and calbindin immunoreactivity in layer 4A.* Cerebral Cortex, 2002. **12**(7): p. 671-691.

34. Preuss, T.M., H. Qi, and J.H. Kaas, *Distinctive compartmental organization of human primary visual cortex.* Proceedings of the National Academy of Sciences of the United States of America, 1999. **96**(20): p. 11601-11606.

35. de Sousa, A.A., et al., *Comparative cytoarchitectural analyses of striate and extrastriate areas in hominoids.* Cerebral Cortex, 2009. **20**(4): p. 966-981.

36. Sherwood, C.C., et al., *Scaling of inhibitory interneurons in areas V1 and V2 of anthropoid primates as revealed by calcium-binding protein immunohistochemistry.* Brain, Behavior and Evolution, 2007. **69**(3): p. 176-195.

37. King, M.-C. and A.C. Wilson, *Evolution at two levels in humans and chimpanzees.* Science, 1975. **188**(4184): p. 107-116.

38. Babbitt, C.C., et al., *Gene expression and adaptive noncoding changes during human evolution.* BMC Genomics, 2017. **18**(1): p. 435.

39. Haygood, R., et al., *Contrasts between adaptive coding and noncoding changes during human evolution.* Proceedings of the National Academy of Sciences of the United States of America, 2010: p. 200911249.

40. Babbitt, C.C., et al., *Both noncoding and protein-coding RNAs contribute to gene expression evolution in the primate brain.* Genome Biology and Evolution, 2010. **2**: p. 67-79.

41. Bauernfeind, A.L., et al., *Evolutionary divergence of gene and protein expression in the brains of humans and chimpanzees.* Genome Biology and Evolution, 2015. **7**(8): p. 2276-2288.

42. Haygood, R., et al., *Promoter regions of many neural-and nutrition-related genes have experienced positive selection during human evolution.* Nature Genetics, 2007. **39**(9): p. 1140.

43. Khaitovich, P., et al., *Parallel patterns of evolution in the genomes and transcriptomes of humans and chimpanzees.* Science, 2005. **309**(5742): p. 1850-1854.

44. Khaitovich, P., et al., *Regional patterns of gene expression in human and chimpanzee brains.* Genome Research, 2004. **14**(8): p. 1462-1473.

45. Brawand, D., et al., *The evolution of gene expression levels in mammalian organs.* Nature, 2011. **478**(7369): p. 343.

46. Blekhman, R., et al., *Gene regulation in primates evolves under tissue-specific selection pressures.* PLoS genetics, 2008. **4**(11): p. e1000271.

47. Sousa, A.M., et al., *Molecular and cellular reorganization of neural circuits in the human lineage.* Science, 2017. **358**(6366): p. 1027-1032.

48. Konopka, G., et al., *Human-specific transcriptional networks in the brain.* Neuron, 2012. **75**(4): p. 601-617.

49. Bernard, A., et al., *Transcriptional architecture of the primate neocortex.* Neuron, 2012. **73**(6): p. 1083-1099.

50. Xu, C., et al., *Human-specific features of spatial gene expression and regulation in eight brain regions.* Genome Research, 2018: p. gr. 231357.117.

51. Oldham, M.C., S. Horvath, and D.H. Geschwind, *Conservation and evolution of gene coexpression networks in human and chimpanzee brains.* Proceedings of the National Academy of Sciences of the United States of America, 2006. **103**(47): p. 17973-17978.

52. Langmead, B. and S.L. Salzberg, *Fast gapped-read alignment with Bowtie 2.* Nature Methods, 2012. **9**(4): p. 357.

53. Kersey, P.J., et al., *Ensembl Genomes 2018: an integrated omics infrastructure for non-vertebrate species.* Nucleic Acids Research, 2017. **46**(D1): p. D802-D808.

54. Schneider, V.A., et al., *Evaluation of GRCh38 and de novo haploid genome assemblies demonstrates the enduring quality of the reference assembly.* Genome Research, 2017.

55. Anders, S., P.T. Pyl, and W. Huber, *HTSeq—a Python framework to work with high-throughput sequencing data.* Bioinformatics, 2015. **31**(2): p. 166-169.

56. Aken, B.L., et al., *Ensembl 2017.* Nucleic Acids Research, 2016. **45**(D1): p. D635-D642.

57. Kinsella, R.J., et al., *Ensembl BioMarts: a hub for data retrieval across taxonomic space.* Database, 2011. **2011**.

58. Robinson, M.D., D.J. McCarthy, and G.K. Smyth, *edgeR: a Bioconductor package for differential expression analysis of digital gene expression data.* Bioinformatics, 2010. **26**(1): p. 139-140.

59. Arnold, C., L.J. Matthews, and C.L. Nunn, *The 10kTrees website: a new online resource for primate phylogeny.* Evolutionary Anthropology, 2010. **19**(3): p. 114-118.

60. Deegan II, J.F. and G.H. Jacobs, *Spectral sensitivity of gibbons: implications for photopigments and color vision.* Folia Primatologica, 2001. **72**(1): p. 26-29.

61. Wheeler, B.C., C.J. Scarry, and A. Koenig, *Rates of agonism among female primates: a cross-taxon perspective.* Behavioral Ecology, 2013. **24**(6): p. 1369-1380.

62. Kamilar, J.M. and N. Cooper, *Phylogenetic signal in primate behaviour, ecology and life history.* Philosophical Transactions of the Royal Society B: Biological Sciences, 2013. **368**(1618): p. 20120341.

63. DeCasien, A.R., S.A. Williams, and J.P. Higham, *Primate brain size is predicted by diet but not sociality.* Nature Ecology & Evolution, 2017. **1**(5): p. 0112.

64. Ogata, H., et al., *KEGG: Kyoto encyclopedia of genes and genomes.* Nucleic Acids Research, 1999. **27**(1): p. 29-34.

65. Antonazzo, G., et al., *Expansion of the Gene Ontology knowledgebase and resources.* Nucleic Acids Research, 2017.

66. Ashburner, M., et al., *Gene Ontology: tool for the unification of biology.* Nature Genetics, 2000. **25**(1): p. 25.

67. Chen, E.Y., et al., *Enrichr: interactive and collaborative HTML5 gene list enrichment analysis tool.* BMC Bioinformatics, 2013. **14**(1): p. 128.

68. Kuleshov, M.V., et al., *Enrichr: a comprehensive gene set enrichment analysis web server 2016 update.* Nucleic Acids Research, 2016. **44**(W1): p. W90-W97.

69. Brown, A.M., R. Wender, and B.R. Ransom, *Metabolic substrates other than glucose support axon function in central white matter.* Journal of Neuroscience Research, 2001. **66**(5): p. 839-843.

70. Tekkök, S.B., et al., *Transfer of glycogen-derived lactate from astrocytes to axons via specific monocarboxylate transporters supports mouse optic nerve activity.* Journal of Neuroscience Research, 2005. **81**(5): p. 644-652.

71. Bauernfeind, A.L. and C.C. Babbitt, *The appropriation of glucose through primate neurodevelopment.* Journal of human evolution, 2014. **77**: p. 132-140.

72. Magistretti, P.J. and I. Allaman, *A cellular perspective on brain energy metabolism and functional imaging.* Neuron, 2015. **86**(4): p. 883-901.

73. Mächler, P., et al., *In vivo evidence for a lactate gradient from astrocytes to neurons.* Cell Metabolism, 2016. **23**(1): p. 94-102.

74. Chen, J., et al., *A quantitative framework for characterizing the evolutionary history of mammalian gene expression.* Genome research, 2019. **29**(1): p. 53-63.

75. Tracey, T.J., et al., *Neuronal Lipid Metabolism: Multiple Pathways Driving Functional Outcomes in Health and Disease.* Frontiers in Molecular Neuroscience, 2018. **11**: p. 10.

76. Vogt, K., *Diversity in GABAergic signaling*, in *Advances in Pharmacology*. 2015, Elsevier. p. 203-222.

77. Fernstrom, J.D. and R.J. Wurtman, *Brain serotonin content: increase following ingestion of carbohydrate diet.* Science, 1971. **174**(4013): p. 1023-1025.

78. Jenkins, T., et al., *Influence of tryptophan and serotonin on mood and cognition with a possible role of the gut-brain axis.* Nutrients, 2016. **8**(1): p. 56.

79. Uddin, M., et al., *Sister grouping of chimpanzees and humans as revealed by genome-wide phylogenetic analysis of brain gene expression profiles.* Proceedings of the National Academy of Sciences of the United States of America, 2004. **101**(9): p. 2957-2962.

80. Stringer, C.B. and P. Andrews, *Genetic and fossil evidence for the origin of modern humans.* Science, 1988. **239**(4845): p. 1263-1268.

81. Schaffner, S.F., et al., *Calibrating a coalescent simulation of human genome sequence variation.* Genome Research, 2005. **15**(11): p. 1576-1583.

82. Fagundes, N.J., et al., *Statistical evaluation of alternative models of human evolution.* Proceedings of the National Academy of Sciences of the United States of America, 2007. **104**(45): p. 17614-17619.

83. Babbitt, C.C., et al., *Genomic signatures of diet-related shifts during human origins.* Proceedings of the Royal Society B: Biological Sciences, 2011. **278**(1708): p. 961-969.

84. Mink, J.W., R.J. Blumenschine, and D.B. Adams, *Ratio of central nervous system to body metabolism in vertebrates: its constancy and functional basis.* American Journal of Physiology-Regulatory, Integrative and Comparative Physiology, 1981. **241**(3): p. R203-R212.

85. Hofman, M.A., *Energy metabolism, brain size and longevity in mammals.* The Quarterly Review of Biology, 1983. **58**(4): p. 495-512.

86. Karbowski, J., *Global and regional brain metabolic scaling and its functional consequences.* BMC Biology, 2007. **5**(1): p. 18.

87. Yu, Y., et al., *Effect of temperature and glia in brain size enlargement and origin of allometric body-brain size scaling in vertebrates.* BMC Evolutionary Biology, 2014. **14**(1): p. 178.

88. Martin, R.D., *Relative brain size and basal metabolic rate in terrestrial vertebrates.* Nature, 1981. **293**(5827): p. 57.

89. Bradley, B.J. and R.R. Lawler, *Linking genotypes, phenotypes, and fitness in wild primate populations.* Evolutionary Anthropology, 2011. **20**(3): p. 104-119.

90. Grinevich, V. and R. Stoop, *Interplay between Oxytocin and Sensory Systems in the Orchestration of Socio-Emotional Behaviors.* Neuron, 2018. **99**(5): p. 887-904.

91. Rilling, J.K., et al., *Intranasal oxytocin modulates neural functional connectivity during human social interaction.* American Journal of Primatology, 2018: p. e22740.

92. Madlon-Kay, S., et al., *Weak effects of common genetic variation in oxytocin and vasopressin receptor genes on rhesus macaque social behavior.* American Journal of Primatology, 2018. **0**(0): p. e22873.

93. Cumming, B.G. and A.J. Parker, *Local disparity not perceived depth is signaled by binocular neurons in cortical area V1 of the macaque.* Journal of Neuroscience, 2000. **20**(12): p. 4758-4767.

94. Gur, M., A. Beylin, and D.M. Snodderly, *Response variability of neurons in primary visual cortex (V1) of alert monkeys.* Journal of Neuroscience, 1997. **17**(8): p. 2914-2920.

95. Supèr, H., H. Spekreijse, and V.A. Lamme, *Two distinct modes of sensory processing observed in monkey primary visual cortex (V1).* Nature Neuroscience, 2001. **4**(3): p. 304.

96. Maddison, W.P. and D.R. Maddison, *Mesquite: a modular system for evolutionary analysis. Version 3.51.* 2018.

97. Varki, A. and T.K. Altheide, *Comparing the human and chimpanzee genomes: searching for needles in a haystack.* Genome research, 2005. **15**(12): p. 1746-1758.

98. Varki, N.M. and A. Varki, *On the apparent rarity of epithelial cancers in captive chimpanzees.* Philosophical Transactions of the Royal Society B: Biological Sciences, 2015. **370**(1673): p. 20140225.

99. Olson, M.V. and A. Varki, *Sequencing the chimpanzee genome: insights into human evolution and disease.* Nature Reviews Genetics, 2003. **4**(1): p. 20.

100. Varki, A., *A chimpanzee genome project is a biomedical imperative.* Genome research, 2000. **10**(8): p. 1065-1070.

101. Shibata, Y., et al., *Extensive evolutionary changes in regulatory element activity during human origins are associated with altered gene expression and positive selection.* PLoS genetics, 2012. **8**(6): p. e1002789.

102. Dvorak, H.F., *Tumors: wounds that do not heal.* New England Journal of Medicine, 1986. **315**(26): p. 1650-1659.

103. Öhlund, D., E. Elyada, and D. Tuveson, *Fibroblast heterogeneity in the cancer wound.* Journal of Experimental Medicine, 2014. **211**(8): p. 1503-1523.

104. Advani, A.S., A.Y. Chen, and C.C. Babbitt, *Human fibroblasts display a differential focal adhesion phenotype relative to chimpanzee.* Evolution, medicine, and public health, 2016. **2016**(1): p. 110-116.

105. Pizzollo, J., et al., *Comparative serum challenges show divergent patterns of gene expression and open chromatin in human and chimpanzee.* Genome biology and evolution, 2018. **10**(3): p. 826-839.

106. Martin, P., *Wound healing--aiming for perfect skin regeneration.* Science, 1997. **276**(5309): p. 75-81.

107. Burridge, K. and M. Chrzanowska-Wodnicka, *Focal adhesions, contractility, and signaling.* Annual review of cell and developmental biology, 1996. **12**(1): p. 463-519.

108. Bershadsky, A.D., et al., *Assembly and mechanosensory function of focal adhesions: experiments and models.* European journal of cell biology, 2006. **85**(3-4): p. 165-173.

109. Tamura, M., et al., *Inhibition of cell migration, spreading, and focal adhesions by tumor suppressor PTEN.* Science, 1998. **280**(5369): p. 1614-1617.

110. Parsons, J.T., *Focal adhesion kinase: the first ten years.* Journal of cell science, 2003. **116**(8): p. 1409-1416.

111. Huttenlocher, A., R.R. Sandborg, and A.F. Horwitz, *Adhesion in cell migration.* Current opinion in cell biology, 1995. **7**(5): p. 697-706.

112. Woods, A. and J.R. Couchman, *Syndecan-4 and focal adhesion function.* Current opinion in cell biology, 2001. **13**(5): p. 578-583.

113. Parsons, J.T., A.R. Horwitz, and M.A. Schwartz, *Cell adhesion: integrating cytoskeletal dynamics and cellular tension.* Nature reviews Molecular cell biology, 2010. **11**(9): p. 633.

114. McLean, G.W., et al., *Specific deletion of focal adhesion kinase suppresses tumor formation and blocks malignant progression.* Genes & development, 2004. **18**(24): p. 2998-3003.

115. Sulzmaier, F.J., C. Jean, and D.D. Schlaepfer, *FAK in cancer: mechanistic findings and clinical applications.* Nature reviews cancer, 2014. **14**(9): p. 598.

116. Tai, Y.-L., L.-C. Chen, and T.-L. Shen, *Emerging roles of focal adhesion kinase in cancer.* BioMed research international, 2015. **2015**.

117. Mygind, K.J., et al., *Loss of ADAM9 expression impairs β1 integrin endocytosis, focal adhesion formation and cancer cell migration.* J Cell Sci, 2018. **131**(1): p. jcs205393.

118. Mao, H., et al., *Hsp72 interacts with paxillin and facilitates the reassembly of focal adhesions during recovery from ATP depletion.* Journal of Biological Chemistry, 2004. **279**(15): p. 15472-15480.

119. Liang, C.-C., A.Y. Park, and J.-L. Guan, *In vitro scratch assay: a convenient and inexpensive method for analysis of cell migration in vitro.* Nature protocols, 2007. **2**(2): p. 329.

120. Nyegaard, S., B. Christensen, and J.T. Rasmussen, *An optimized method for accurate quantification of cell migration using human small intestine cells.* Metabolic Engineering Communications, 2016. **3**: p. 76-83.

121. Schindelin, J., et al., *Fiji: an open-source platform for biological-image analysis.* Nature methods, 2012. **9**(7): p. 676.

122. O'Leary, C.J., M. Weston, and K.W. McDermott, *An ex vivo model to quantitatively analyze cell migration in tissue.* Developmental Dynamics, 2018. **247**(1): p. 201-211.

123. Maiuri, P., et al., *The first World Cell Race.* Current biology : CB, 2012. **22**(17): p. R673-R675.

124. Lamprecht, M.R., D.M. Sabatini, and A.E. Carpenter, *CellProfiler™: free, versatile software for automated biological image analysis.* Biotechniques, 2007. **42**(1): p. 71-75.

125. Gupton, S.L. and C.M. Waterman-Storer, *Spatiotemporal feedback between actomyosin and focal-adhesion systems optimizes rapid cell migration.* Cell, 2006. **125**(7): p. 1361-1374.

126. Wehrle-Haller, B., *Structure and function of focal adhesions.* Current opinion in cell biology, 2012. **24**(1): p. 116-124.

127. McClure, H.M., *Tumors in nonhuman primates: Observations during a six-year period in the Yerkes Primate center colony.* American journal of physical anthropology, 1973. **38**(2): p. 425-429.

128. Seibold, H., *Neoplasms and proliferative lesions in 1065 nonhuman primate necropsies.* Lab Anim Sci, 1973. **23**: p. 533-537.

129. Beniashvili, D.S., *An overview of the world literature on spontaneous tumors in nonhuman primates.* Journal of medical primatology, 1989. **18**(6): p. 423-437.

130. Puente, X.S., et al., *Comparative analysis of cancer genes in the human and chimpanzee genomes.* BMC genomics, 2006. **7**(1): p. 15.

131. Sullivan, T.P., et al., *The pig as a model for human wound healing.* Wound repair and regeneration, 2001. **9**(2): p. 66-76.

132. Chang, H.Y., et al., *Robustness, scalability, and integration of a wound-response gene expression signature in predicting breast cancer survival.* Proceedings of the National Academy of Sciences, 2005. **102**(10): p. 3738-3743.

133. Chang, H.Y., et al., *Gene expression signature of fibroblast serum response predicts human cancer progression: similarities between tumors and wounds.* PLoS biology, 2004. **2**(2): p. e7.

134. Desgrosellier, J.S. and D.A. Cheresh, *Integrins in cancer: biological implications and therapeutic opportunities.* Nature Reviews Cancer, 2010. **10**(1): p. 9.

135. Ou, J., et al., *Endostatin suppresses colorectal tumor-induced lymphangiogenesis by inhibiting expression of fibronectin extra domain a and integrin α9.* Journal of cellular biochemistry, 2011. **112**(8): p. 2106-2114.

136. Ming, X.-Y., et al., *Integrin α7 is a functional cancer stem cell surface marker in oesophageal squamous cell carcinoma.* Nature communications, 2016. **7**: p. 13568.

137. Haas, T.L., et al., *Integrin α7 is a functional marker and potential therapeutic target in glioblastoma.* Cell Stem Cell, 2017. **21**(1): p. 35-50. e9.

138. Zhang, Y.-L., et al., *Integrin α9 Suppresses Hepatocellular Carcinoma Metastasis by Rho GTPase Signaling.* Journal of Immunology Research, 2018. **2018**.

139. Laszlo, V., et al., *Epigenetic down-regulation of integrin α7 increases migratory potential and confers poor prognosis in malignant pleural mesothelioma.* The Journal of pathology, 2015. **237**(2): p. 203-214.

140. Xian, X., S. Gopal, and J.R. Couchman, *Syndecans as receptors and organizers of the extracellular matrix.* Cell and tissue research, 2010. **339**(1): p. 31.

141. Varki, A. and D.L. Nelson, *Genomic comparisons of humans and chimpanzees.* Annu. Rev. Anthropol., 2007. **36**: p. 191-209.

142. Kuzawa, C.W., et al., *Metabolic costs and evolutionary implications of human brain development.* Proceedings of the National Academy of Sciences, 2014. **111**(36): p. 13010-13015.

143. Herculano-Houzel, S., *Scaling of brain metabolism with a fixed energy budget per neuron: implications for neuronal activity, plasticity and evolution.* PloS one, 2011. **6**(3): p. e17514.

144. Aiello, L.C. and P. Wheeler, *The expensive-tissue hypothesis: the brain and the digestive system in human and primate evolution.* Current Anthropology, 1995. **36**(2): p. 199-221.

145. Brown, F., et al., *Early Homo erectus skeleton from west lake Turkana, Kenya.* Nature, 1985. **316**(6031): p. 788.

146. McHenry, H.M., *Body size and proportions in early hominids.* American Journal of Physical Anthropology, 1992. **87**(4): p. 407-431.

147. McHenry, H.M., *Tempo and mode in human evolution.* Proceedings of the National Academy of Sciences, 1994. **91**(15): p. 6780-6786.

148. Peters, C.R., *Theoretical and actualistic ecobotanical perspectives on early hominin diets and paleoecology.* Evolution of the human diet: the known, the unknown, and the unknowable, 2007: p. 233-261.

149. Shea, J.J., *Lithic archaeology, or, what stone tools can (and can't) tell us about early hominin diets.* Evolution of the human diet: the known, the unknown and the unknowable, 2007: p. 321-351.

150. West, G.B., J.H. Brown, and B.J. Enquist, *A general model for ontogenetic growth.* Nature, 2001. **413**(6856): p. 628.

151. Pontzer, H., et al., *Primate energy expenditure and life history.* Proceedings of the National Academy of Sciences of the United States of America, 2014. **111**(4): p. 1433-1437.

152. Stearns, S., *The evolution of life histories*. 1992, Oxford Univ. Press.

153. Leonard, W.R. and M.L. Robertson, *Comparative primate energetics and hominid evolution.* American Journal of Physical Anthropology: The Official Publication of the American Association of Physical Anthropologists, 1997. **102**(2): p. 265-281.

154. Leonard, W.R. and M.L. Robertson, *Evolutionary perspectives on human nutrition: the influence of brain and body size on diet and metabolism.* American Journal of Human Biology, 1994. **6**(1): p. 77-88.

155. Leonard, W.R., et al., *Metabolic correlates of hominid brain evolution.* Comparative Biochemistry and Physiology Part A: Molecular & Integrative Physiology, 2003. **136**(1): p. 5-15.

156. Pontzer, H., et al., *Constrained total energy expenditure and metabolic adaptation to physical activity in adult humans.* Current Biology, 2016. **26**(3): p. 410-417.

157. Nelson, D.L., A.L. Lehninger, and M.M. Cox, *Lehninger principles of biochemistry*. 2008: Macmillan.

158. Raichle, M.E., *Two views of brain function.* Trends in cognitive sciences, 2010. **14**(4): p. 180-190.

159. Vander Heiden, M.G., L.C. Cantley, and C.B. Thompson, *Understanding the Warburg effect: the metabolic requirements of cell proliferation.* science, 2009. **324**(5930): p. 1029-1033.

160. Vander Heiden, M.G., et al., *Evidence for an alternative glycolytic pathway in rapidly proliferating cells.* Science, 2010. **329**(5998): p. 1492-1499.

161. Kosiol, C., et al., *Patterns of positive selection in six mammalian genomes.* PLoS genetics, 2008. **4**(8): p. e1000144.

162. Uddin, M., et al., *Distinct genomic signatures of adaptation in pre-and postnatal environments during human evolution.* Proceedings of the National Academy of Sciences, 2008. **105**(9): p. 3215-3220.

163. Pollard, K.S., et al., *Forces shaping the fastest evolving regions in the human genome.* PLoS genetics, 2006. **2**(10): p. e168.

164. Prabhakar, S., et al., *Accelerated evolution of conserved noncoding sequences in humans.* Science, 2006. **314**(5800): p. 786-786.

165. Preuss, T.M., *Human brain evolution: from gene discovery to phenotype discovery.* Proceedings of the National Academy of Sciences, 2012. **109**(Supplement 1): p. 10709-10716.

166. Goldberg, A., et al., *Adaptive evolution of cytochrome c oxidase subunit VIII in anthropoid primates.* Proceedings of the National Academy of Sciences of the United States of America, 2003. **100**(10): p. 5873-5878.

167. Grossman, L.I., et al., *Molecular evolution of aerobic energy metabolism in primates.* Molecular phylogenetics and evolution, 2001. **18**(1): p. 26-36.

168. Grossman, L.I., et al., *Accelerated evolution of the electron transport chain in anthropoid primates.* TRENDS in Genetics, 2004. **20**(11): p. 578-585.

169. Hüttemann, M., et al., *Regulation of mitochondrial respiration and apoptosis through cell signaling: cytochrome c oxidase and cytochrome c in ischemia/reperfusion injury and inflammation.* Biochimica et Biophysica Acta (BBA)-Bioenergetics, 2012. **1817**(4): p. 598-609.

170. Romero, I.G., et al., *A panel of induced pluripotent stem cells from chimpanzees: a resource for comparative functional genomics.* Elife, 2015. **4**: p. e07103.

171. Almeida, A., S. Moncada, and J.P. Bolaños, *Nitric oxide switches on glycolysis through the AMP protein kinase and 6-phosphofructo-2-kinase pathway.* Nature cell biology, 2004. **6**(1): p. 45.

172. Herrero-Mendez, A., et al., *The bioenergetic and antioxidant status of neurons is controlled by continuous degradation of a key glycolytic enzyme by APC/C–Cdh1.* Nature cell biology, 2009. **11**(6): p. 747.

173. Sonntag, K.-C., et al., *Late-onset Alzheimer's disease is associated with inherent changes in bioenergetics profiles.* Scientific reports, 2017. **7**(1): p. 14038.

174. Nedergaard, M., B. Ransom, and S.A. Goldman, *New roles for astrocytes: redefining the functional architecture of the brain.* Trends in neurosciences, 2003. **26**(10): p. 523-530.

175. Pellerin, L. and P.J. Magistretti, *Glutamate uptake into astrocytes stimulates aerobic glycolysis: a mechanism coupling neuronal activity to glucose utilization.* Proceedings of the National Academy of Sciences, 1994. **91**(22): p. 10625-10629.

176. Volkenhoff, A., et al., *Glial glycolysis is essential for neuronal survival in Drosophila.* Cell Metabolism, 2015. **22**(3): p. 437-447.

177. Diniz, L.P., et al., *Astrocyte-induced synaptogenesis is mediated by transforming growth factor β signaling through modulation of D-serine levels in cerebral cortex neurons.* Journal of Biological Chemistry, 2012. **287**(49): p. 41432-41445.

178. Meyer-Franke, A., et al., *Characterization of the signaling interactions that promote the survival and growth of developing retinal ganglion cells in culture.* Neuron, 1995. **15**(4): p. 805-819.

179. Blake, L.E., et al., *A comparative study of endoderm differentiation in humans and chimpanzees.* Genome biology, 2018. **19**(1): p. 162.

180. Burrows, C.K., et al., *Genetic variation, not cell type of origin, underlies the majority of identifiable regulatory differences in iPSCs.* PLoS genetics, 2016. **12**(1): p. e1005793.

181. Eres, I.E., et al., *Reorganization of 3D genome structure may contribute to gene regulatory evolution in primates.* PLoS genetics, 2019. **15**(7): p. e1008278.

182. Pavlovic, B.J., et al., *A comparative assessment of human and chimpanzee iPSC-derived cardiomyocytes with primary heart tissues.* Scientific reports, 2018. **8**(1): p. 15312.

183. Ward, M.C. and Y. Gilad, *A generally conserved response to hypoxia in iPSC-derived cardiomyocytes from humans and chimpanzees.* eLife, 2019. **8**: p. e42374.

184. Ward, M.C., et al., *Silencing of transposable elements may not be a major driver of regulatory evolution in primate iPSCs.* Elife, 2018. **7**: p. e33084.

185. Edgar, R., M. Domrachev, and A.E. Lash, *Gene Expression Omnibus: NCBI gene expression and hybridization array data repository.* Nucleic acids research, 2002. **30**(1): p. 207-210.

186. Zhang, Y., et al., *Purification and characterization of progenitor and mature human astrocytes reveals transcriptional and functional differences with mouse.* Neuron, 2016. **89**(1): p. 37-53.

187. Consortium, E.P., *An integrated encyclopedia of DNA elements in the human genome.* Nature, 2012. **489**(7414): p. 57-74.

188. Davis, C.A., et al., *The Encyclopedia of DNA elements (ENCODE): data portal update.* Nucleic acids research, 2018. **46**(D1): p. D794-D801.

189. Raudvere, U., et al., *g:Profiler: a web server for functional enrichment analysis and conversions of gene lists (2019 update).* Nucleic Acids Research, 2019. **47**(W1): p. W191-W198.

190. Subramanian, A., et al., *Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles.* Proceedings of the National Academy of Sciences, 2005. **102**(43): p. 15545-15550.

191. Liberzon, A., et al., *Molecular signatures database (MSigDB) 3.0.* Bioinformatics, 2011. **27**(12): p. 1739-1740.

192. Herrero, J., et al., *Ensembl comparative genomics resources.* Database, 2016. **2016**.

193. Romero, P., et al., *Computational prediction of human metabolic pathways from the complete human genome.* Genome biology, 2005. **6**(1): p. R2.

194. McKenzie, A.T., et al., *Brain Cell Type Specific Gene Expression and Co-expression Network Architectures.* Scientific Reports, 2018. **8**(1): p. 8868.

195. Fabregat, A., et al., *The Reactome Pathway Knowledgebase.* Nucleic acids research, 2018. **46**(D1): p. D649-D655.

196. Reimand, J., et al., *Pathway enrichment analysis and visualization of omics data using g:Profiler, GSEA, Cytoscape and EnrichmentMap.* Nature protocols, 2019. **14**(2): p. 482-517.

197. Maxson, M.E. and S. Grinstein, *The vacuolar-type H+-ATPase at a glance– more than a proton pump*. 2014, The Company of Biologists Ltd.

198. Pamarthy, S., et al., *The curious case of vacuolar ATPase: regulation of signaling pathways.* Molecular cancer, 2018. **17**(1): p. 41.

199. Almad, A.A., et al., *Connexin 43 in astrocytes contributes to motor neuron toxicity in amyotrophic lateral sclerosis.* Glia, 2016. **64**(7): p. 1154-1169.

200. Yang, Z., *PAML 4: phylogenetic analysis by maximum likelihood.* Molecular biology and evolution, 2007. **24**(8): p. 1586-1591.

201. Horvath, J.E., et al., *Genetic comparisons yield insight into the evolution of enamel thickness during human evolution.* Journal of human evolution, 2014. **73**: p. 75-87.

202. Muntané, G., et al., *Analysis of synaptic gene expression in the neocortex of primates reveals evolutionary changes in glutamatergic neurotransmission.* Cerebral Cortex, 2014. **25**(6): p. 1596-1607.

203. Pond, S.L.K., S.D.W. Frost, and S.V. Muse, *HyPhy: hypothesis testing using phylogenies.* Bioinformatics, 2004. **21**(5): p. 676-679.

204. Charnov, E.L. and D. Berrigan, *Why do female primates have such long lifespans and so few babies? Or life in the slow lane.* Evolutionary Anthropology, 1993. **1**(6): p. 191-194.

205. Snodgrass, J.J., W.R. Leonard, and M.L. Robertson, *Primate bioenergetics: an evolutionary perspective*, in *Primate origins: adaptations and evolution*. 2007, Springer. p. 703-737.

206. Lovegrove, B.G., *Age at first reproduction and growth rate are independent of basal metabolic rate in mammals.* Journal of Comparative Physiology B, 2009. **179**(4): p. 391.

207. Speakman, J.R., *Body size, energy metabolism and lifespan.* Journal of Experimental Biology, 2005. **208**(9): p. 1717-1730.